线性代数原理

张远达 编

上海教育出版社

线性代数原理

张远达 编

上海教育出版社出版

（上海永福路 123 号）

新华书店上海发行所发行　上海群众印刷厂印刷

开本 787×1092 1/32　印张 16.75　字数 366,000

1980 年 8 月第 1 版　1980 年 8 月第 1 次印刷

印数 1—20,000 本

统一书号：7150·2185　　定价：1.85 元

前　言

线性代数这门学科，在十九世纪就已经获得了光辉的成就. 由于它在数学的许多分支以及物理、技术科学中有越来越广泛的应用，所以直到目前，它不仅在代数，而且在科学技术的许多领域里，都占有重要的地位. 正因为这个原故，需要学习和了解线性代数的读者越来越多. 作为一个数学工作者，写一本便于自修线性代数的书籍，应为刻不容缓的.

编写这书的目的，是想详细地、系统地阐述线性代数的主要内容，以利于读者自修. 因此，本书在写法上，每章甚至每节的开头，总是把要解决的问题明确提出，让读者先有一个梗概. 对于线性代数的初等部分，如行列式和线性方程组的理论，仍给以一定的篇幅. 对于非齐次(线性的)方程组的基础解系这个问题，在一般书上就谈得很少，而本书完全用讲齐次方程组之基础解系的平行的办法来讲，借以对照看出齐次与非齐次之间无实质性的差异. 而讲基础解系时势必牵涉到向量间的线性相关及无关的概念，这对初学者理解时往往感觉困难较大，本书环绕"行列式等于零的充要条件是它的行(或列)为线性相关"这个问题为中心来写是合宜的.

线性代数传统的写法大致有两种，一是从线性空间的线性变换的所谓不变子空间的概念出发，引出线性变换的一些标准形(如有理标准形、约旦标准形、对角形等)；另一种写法是从矩阵的角度入手，直接讲求矩阵的这样一些标准形的具体方法. 前一种写法的优点在于把几何问题能够阐明得非常

透彻，但其缺点是对某一个矩阵怎样用具体的方法求得它的标准形，往往使初学的同志感到陌生；后一种写法固可使读者明了怎样实际求矩阵之标准形的方法，但对于其几何意义则茫无所知．线性代数实际上产生于解析几何，线性代数与解析几何有紧密的联系，故若丢掉几何意义，只讲矩阵之标准形的实际求法，这不是一个好的方法．然考虑到矩阵是数学本身以及应用科学的一个很重要的工具，又不得不突出矩阵，所以本书是突出矩阵的运算，但紧接着就详尽阐述它们的几何意义，借以开阔读者的视野，洞悉几何与代数之内在的联系．

因此，全书自始至终，对代数与几何的关系都给以重视，例如坐标变换、二次曲面（曲线）的分类（投影、仿射、度量）都反复作了说明，这样既联系了几何，也丰富了线代数的内容．同样，在讲了实直交矩阵之实标准形后（第八章），也跟着说明直觉空间（三维与二维的）内的直交变换究竟表示怎样的几何运动．

当然，由于本书是研究线性代数，所以对于提及的几何材料不能要求自成几何系统，只能把几何材料看成是线性代数的应用．线性代数的应用自然不只限于解析几何，如数学分析和微分方程，它们与线性代数的关系也很密切，因而在第六章末说明了怎样利用二次齐式中有定及不定的理论来解决数学分析中多元函数之极大极小问题，在第五章第6节后面紧接着又介绍了利用矩阵之有理标准形与约旦标准形来解微分方程组．本书尤其对线性代数在微分方程中的应用给以特别的重视，故第九章讲矩阵分析，目的是说明怎样利用矩阵知识来解决一阶变系数线性非齐次微分方程组的解之存在及唯一性问题．同时，也在第九章，利用矩阵分析知识研究了在线性代数计算方法中解线性方程组之迭代法的收敛条件．

第十章讨论几个特殊问题，如直积、复合矩阵、特征值的估计、矩阵之极分解、全矩阵环的基. 利用复合矩阵而通过极分解，说明酉矩阵与厄米特矩阵在线性代数中的重要地位，因而对酉矩阵与厄米特矩阵的特性又作了深入一步的探讨.

1962 年时，作者曾与熊全淹先生合编《线性代数》一书，近年来，学习(包括自修)线性代数的读者渐多，各方来信询问这书甚为殷切，但考虑作者在 62 年的书对自学者不尽相宜，鉴于各方面的这一要求理应重新改写，限于个人学识水平，不妥甚至谬误之处一定很多，望广大读者指正.

武汉大学　张远达

一九七九年三月

目　　录

第一章　行　列　式

本章讨论下面五个问题：1．行列式的概念是怎样形成的？2．行列式有什么性质？3．怎样计算一个行列式？4．如何应用行列式的性质来解线性方程组？5．行列式的性质与计算可以怎样推广？

第1节　行列式的概念

行列式这个概念究竟是怎样形成的呢？这就得从解一次方程组谈起．

我们都知道怎样解方程组：

$$\begin{cases} a_1x+b_1y=k_1, \\ a_2x+b_2y=k_2. \end{cases} \tag{1}$$

具体步骤是：从(1)里先消去 y 而求得 x，这只要将(1)的第一、二两个式子分别乘以 b_2 与 $-b_1$，然后再相加，就得到

$$(a_1b_2-a_2b_1)x=k_1b_2-k_2b_1.$$

同理，在(1)中也可消去 x 而求得 y，这就要将(1)的第一、二两个式子分别乘以 $-a_2$ 与 a_1，然后再相加，得到

$$(a_1b_2-a_2b_1)y=a_1k_2-a_2k_1.$$

于是，在 $D=a_1b_2-a_2b_1\neq0$ 时，(1)的解为

$$\begin{cases} x=\dfrac{k_1b_2-k_2b_1}{a_1b_2-a_2b_1}, \\ y=\dfrac{a_1k_2-a_2k_1}{a_1b_2-a_2b_1}. \end{cases} \tag{2}$$

反之，容易检验(2)又确为(1)的解.

(2)式尽管提供了方程组(1)之解的公式，但(2)式很难记忆. 为了易于记住(2)式，我们要从表示式中发现一些字母之间的规律.

可以发现，(2)中 x 与 y 的表示式中的分母是相同的，即都等于 $a_1b_2-a_2b_1$，而这个数又仅与(1)中 x, y 的系数有关，与常数项无关. 如果把这些系数按它们在原来方程组中的位置写出，并在两旁各加一条竖直线，即用符号

$$\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix}$$

来表示 $a_1b_2-a_2b_1$ 这个数，即

$$D=\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix}=a_1b_2-a_2b_1. \tag{3}$$

我们把(3)称为*二级行列式*，把 D 中横写的叫做*行*，竖写的叫做*列*，因此(3)含有两个行与两个列. 从(3)式看出：符号 $\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix}$ 可看成是：从左上角到右下角的对角线上两个元素相乘取符号"+"所得的 a_1b_2，再将其中从右上角到左下角的对角线上两个元素相乘取符号"−"所得的 $-a_2b_1$，然后再把这两个结果合并就得到了(3)式. 这就是计算二级行列式的规则.

利用这个规则，即可把(2)中 x 与 y 之表达式的分子分别表成

$$\begin{vmatrix} k_1 & b_1 \\ k_2 & b_2 \end{vmatrix} \quad 与 \quad \begin{vmatrix} a_1 & k_1 \\ a_2 & k_2 \end{vmatrix},$$

于是(2)式可记成

$$x=\frac{\begin{vmatrix} k_1 & b_1 \\ k_2 & b_2 \end{vmatrix}}{\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix}}, \quad y=\frac{\begin{vmatrix} a_1 & k_1 \\ a_2 & k_2 \end{vmatrix}}{\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix}}.$$

象这样用二级行列式来表达方程组(1)的解的公式，非常简便，而且也容易记忆.

再看含三个未知量 x_1, x_2, x_3 的三个方程所成的方程组

$$\begin{cases} a_{11}x_1+a_{12}x_2+a_{13}x_3=b_1, \\ a_{21}x_1+a_{22}x_2+a_{23}x_3=b_2, \\ a_{31}x_1+a_{32}x_2+a_{33}x_3=b_3. \end{cases} \tag{4}$$

为了解(4)式，就得消去(4)中的任两个未知量，求得只含一个未知量的方程. 例如把(4)的第一、二、三个方程分别用 $a_{22}a_{33}-a_{23}a_{32}$, $a_{13}a_{32}-a_{12}a_{33}$, $a_{12}a_{23}-a_{13}a_{22}$ 去乘，然后再加起来，结果可知 x_2 与 x_3 的系数都等于零，而得到等式：

$$\begin{aligned}&(a_{11}a_{22}a_{33}+a_{12}a_{23}a_{31}+a_{13}a_{21}a_{32}-a_{11}a_{23}a_{32}-a_{12}a_{21}a_{33}\\ &\quad -a_{13}a_{22}a_{31})x_1\\ &=b_1a_{22}a_{33}+a_{12}a_{23}b_3+a_{13}b_2a_{32}-b_1a_{23}a_{32}\\ &\quad -a_{12}b_2a_{33}-a_{13}a_{22}b_3.\end{aligned}$$

所以，当

$$\begin{aligned}D=&a_{11}a_{22}a_{33}+a_{12}a_{23}a_{31}+a_{13}a_{21}a_{32}-a_{11}a_{23}a_{32}\\ &-a_{12}a_{21}a_{33}-a_{13}a_{22}a_{31}\neq 0\end{aligned}$$

时，就得到

$$x_1=\frac{b_1a_{22}a_{33}+a_{12}a_{23}b_3+a_{13}b_2a_{32}-b_1a_{23}a_{32}-a_{12}b_2a_{33}-a_{13}a_{22}b_3}{a_{11}a_{22}a_{33}+a_{12}a_{23}a_{31}+a_{13}a_{21}a_{32}-a_{11}a_{23}a_{32}-a_{12}a_{21}a_{33}-a_{13}a_{22}a_{31}}.$$

用类似的方法，可求得

$$x_2=\frac{a_{11}b_2a_{33}+b_1a_{23}a_{31}+a_{13}a_{21}b_3-a_{11}a_{23}b_3-b_1a_{21}a_{33}-a_{13}b_2a_{31}}{a_{11}a_{22}a_{33}+a_{12}a_{23}a_{31}+a_{13}a_{21}a_{32}-a_{11}a_{23}a_{32}-a_{12}a_{21}a_{33}-a_{13}a_{22}a_{31}},$$

$$x_3=\frac{a_{11}a_{22}b_3+a_{12}b_2a_{31}+b_1a_{21}a_{32}-a_{11}b_2a_{32}-a_{12}a_{21}b_3-b_1a_{22}a_{31}}{a_{11}a_{22}a_{33}+a_{12}a_{23}a_{31}+a_{13}a_{21}a_{32}-a_{11}a_{23}a_{32}-a_{12}a_{21}a_{33}-a_{13}a_{22}a_{31}}.$$

这就是说：当 $D\neq 0$ 时，方程组(4)如果有解，这解必是上面那样的形状．

反之，将这样的 x_1, x_2, x_3 的值代入方程组(4)，也容易检验它们确是(4)的解．所以，当 $D\neq 0$ 时，方程组(4)有唯一的一组解，即是上面 x_1, x_2, x_3 的表达式．

同前面一样，为了便于记忆这公式，我们要从这些表达式中发现一些字母之间的规律．

首先，我们看到 x_1, x_2, x_3 的表达式里分母都等于 D；而 $x_i(i=1, 2, 3)$ 的表达式之分子，是将 D 中凡 x_i 之系数 a_{1i}, a_{2i}, a_{3i} 分别用 b_1, b_2, b_3 去替换而成的．于是，只要对分母的结构弄清楚了，则 x_1, x_2, x_3 之表达式的结构也就搞清楚了．

与引进二级行列式一样，用符号

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} \tag{5}$$

来表示 D，即

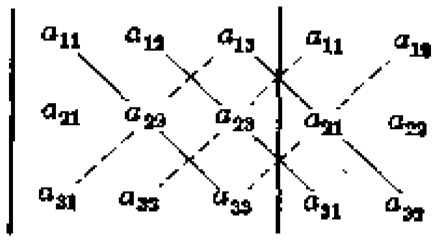
$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11}a_{22}a_{33}+a_{12}a_{23}a_{31}+a_{13}a_{21}a_{32}-a_{11}a_{23}a_{32}-a_{12}a_{21}a_{33}-a_{13}a_{22}a_{31};$$

(5)中含有三行、三列，我们把(5)称为**三级行列式**．上式表明，(5)是由六个项组成，其中每一项都是(5)中三个元素的乘积，且每个乘积含有每一行及每一列的一个元素且仅含每一行每一列的一个元素，因此，其中的每一项都可以写成下面的形式：

$$a_{1p}a_{2q}a_{3r}, \tag{6}$$

其中 p, q, r 是自然数 1, 2, 3 的某排列．与二级行列式类

似，引用对角线法则，很容易把(5)中正负项写出来．先写出一个有六条对角线的表，如



这对角线表是将原三级行列式的第一、二两列重复写出依次附加在三级行列式的右边而成，其中每一条对角线把三个元素连在一块取其积，在实线对角线上三个元素的乘积冠以“+”号，在虚线对角线上三个元素的乘积冠以“−”号，然后把这六个乘积加起来，就得到三级行列式(5)的值．

引用三级行列式的记号，对方程组(4)，当 $D\neq 0$ 时，它的唯一的一组解就可写成：

$$x_1=\frac{\begin{vmatrix} b_1 & a_{12} & a_{13}\\ b_2 & a_{22} & a_{23}\\ b_3 & a_{32} & a_{33}\end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} & a_{13}\\ a_{21} & a_{22} & a_{23}\\ a_{31} & a_{32} & a_{33}\end{vmatrix}},\qquad x_2=\frac{\begin{vmatrix} a_{11} & b_1 & a_{13}\\ a_{21} & b_2 & a_{23}\\ a_{31} & b_3 & a_{33}\end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} & a_{13}\\ a_{21} & a_{22} & a_{23}\\ a_{31} & a_{32} & a_{33}\end{vmatrix}},$$

$$x_3=\frac{\begin{vmatrix} a_{11} & a_{12} & b_1\\ a_{21} & a_{22} & b_2\\ a_{31} & a_{32} & b_3\end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} & a_{13}\\ a_{21} & a_{22} & a_{23}\\ a_{31} & a_{32} & a_{33}\end{vmatrix}}.$$

怎样把二级、三级行列式的意义推广到一般的 n 级行列式，并利用它来表达由 n 个未知量 n 个方程所组成的线性方程组的

解呢？从上面的推导可知，由两个未知量中消去一个比较容易，而由三个未知量中消去两个就已经很麻烦了，至于一般，由 n 个未知量中消去 $n-1$ 个，在理论上虽然是可能的，但要具体地实施，其困难程度可想而知．因此，n 级行列式的定义也就不能用上面类似的方法去求得．

我们来详细研究一下二级、三级行列式的结构，找出它们内在的规律，然后根据它来形式地定义 n 级行列式．可以看出，为了得到(5)中所有的项，只需要在乘积(6)中使第二个指标 $p,\ q,\ r$ 取 1, 2, 3 的所有可能的六种不同的排列

$$1,\ 2,\ 3;\ 2,\ 3,\ 1;\ 3,\ 1,\ 2;\ 1,\ 3,\ 2;\ 2,\ 1,\ 3;\ 3,\ 2,\ 1 \qquad (7)$$

就行了．但是，乘积(6)在(5)中出现的时候，有一些带有正号，另一些却带有负号，这就还需要指明选择正负号时所依据的法则．实际上，(6)中带有正号的那些项(即乘积)的第二个指标 $p,\ q,\ r$ 形成下列的三个排列：

$$1,\ 2,\ 3; \qquad 2,\ 3,\ 1; \qquad 3,\ 1,\ 2; \qquad (8)$$

带有负号的那些乘积的第二个指标 $p,\ q,\ r$ 形成下列的三个排列：

$$1,\ 3,\ 2; \qquad 2,\ 1,\ 3; \qquad 3,\ 2,\ 1. \qquad (9)$$

现在问(8)中的三个排列与(9)中的三个排列的区别是什么？为此，先要引入一个逆序的概念．

任意两个自然数，如果大的在小的前面，就叫这两个自然数间有一个*逆序*．一个排列中所有任意两个数的逆序数的和，叫做这排列的*逆序数*．在(8)中，对第一个排列，没有逆序(即逆序数等于零)；在第二个排列中，如果逐次把每一个数同它后面的各个数比较，容易看出有两个逆序(一个是2在1前，另一个是3在1前)；同样，可以看出，第三个排列也有两个逆序(一个是3在1前，另一个是3在2前)．于是，我们得知

(8)中所有的排列都有偶数个逆序．用同样的方法可知，(9)中所有的排列都有奇数个逆序．当排列的逆序数为偶数或奇数时，就叫该排列分别为偶排列或奇排列．这就告诉我们，(5)中各项正负号选择的法则是：(5)中任一项(6)应冠以正号或负号，完全由它的第二个指标形成的排列 p, q, r 为偶排列或奇排列来决定．

在上述讨论基础上，我们可以得到：

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = \sum_{(p,q,r)} (-1)^{[p,q,r]} a_{1p}a_{2q}a_{3r}, \tag{10}$$

式中符号 $[p, q, r]$ 表示排列 p, q, r 的逆序数，符号 $\sum\limits_{(p,q,r)}$ 表示对 1, 2, 3 的所有可能的排列取和．

因为 1, 2 两个数的所有可能的排列只有 1, 2 与 2, 1 两个，前者的逆序数为零，后者的逆序数为 1，所以

$$\sum_{(p,q)} (-1)^{[p,q]} a_{1p}a_{2q} = (-1)^{[1,2]} a_{11}a_{22} + (-1)^{[2,1]} a_{12}a_{21}$$

$$= a_{11}a_{22} - a_{12}a_{21},$$

因之，有

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = \sum_{(p,q)} (-1)^{[p,q]} a_{1p}a_{2q},$$

这恰与(10)式的形式一致．这就是说，二级行列式的概念为三级行列式的特例．在(10)式基础上，不难推广到一般的情形：设有 n^2 个数，把它们写成一个有 n 行和 n 列的符号，如

$$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}. \tag{11}$$

我们叫(11)为 n 级行列式，它表示所有这样乘积的代数和，即

每个乘积含有 n 个元素，使每行每列的元有且只有一个出现在该乘积中，于是这些乘积都可以写成下面的形式：

$$a_{1p_1}a_{2p_2}\cdots a_{np_n}, \tag{12}$$

其中 $p_1, p_2, \cdots, p_n$ 是 1, 2, $\cdots$, n 的某一排列，每个这样的排列依 1, 2, $\cdots$, n 的顺序来比较就有一个逆序数，当这逆序数为偶数时，这项取(+)号，而当逆序数为奇数时，这项就取负(−)号．用式子表示，就是：

$$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}$$

$$= \sum_{(p_1, p_2, \cdots, p_n)} (-1)^{[p_1, p_2, \cdots, p_n]} \cdot a_{1p_1}a_{2p_2}\cdots a_{np_n}, \tag{13}$$

其中符号 $[p_1, p_2, \cdots, p_n]$ 表示排列 $p_1, p_2, \cdots, p_n$ 的逆序数，符号 $\sum\limits_{(p_1, p_2, \cdots, p_n)}$ 表示要对 1, 2, $\cdots$, n 的所有排列取和．

因为 n 个数字 1, 2, $\cdots$, n 的所有排列共有 $1\cdot 2\cdot\cdots\cdot n = n!$ 个，所以 n 级行列式共有 $n!$ 个项．

还需说明的是，对于行列式(11)中的每个元素 a_{ij}，它的右下角有两个下标，第一个下标"i"表示元素 a_{ij} 所在的行的序数，第二个下标"j"则表示这元素所在的列的序数，于是 a_{ij} 又表明它是第 i 行第 j 列的元素．面(13)中把每项写成乘积 $a_{1p_1}a_{2p_2}\cdots a_{np_n}$ 的形式，其意义就是要将每项的 n 个元素依照行的先后次序来写(例如 a_{1p_1} 是在第一行上的，所以写在乘积的最前面，a_{2p_2} 是在第二行上的，所以写在乘积的第二位，最后 a_{np_n} 是在第 n 行上的，所以写在乘积的第 n 位)．所谓 $p_1, p_2, \cdots, p_n$ 依顺序 1, 2, $\cdots$, n 来比较其逆序数，就是说按照列的顺序来比较逆序数．至于当元素 a_{ij} 的第一个下标 i 不

是表示它所在的行的序数，或第二个下标 j 不是表示它所在的列的序数时，表示式(13)中的任意项 $a_{1p_1}a_{2p_2}\cdots a_{np_n}$ 的排法以及作为与 $p_1, p_2, \cdots, p_n$ 比较的顺序 1, 2, …, n 都应当相应地改变. 例如

$$\begin{vmatrix} a_{23} & a_{22} & a_{21} \\ a_{13} & a_{12} & a_{11} \\ a_{33} & a_{32} & a_{31} \end{vmatrix} = \sum_{(p,q,r)} (-1)^{[p,q,r]} a_{2p}a_{1q}a_{3r},$$

这里 $[p, q, r]$ 是排列 (p, q, r) 依顺序 3, 2, 1 比较所得的逆序数. 因为第一、二、三列是以右下角的第二个下标 3、2、1 这三个数分别来表示的，所以有 $[3, 2, 1]=0$, $[1, 3, 2]=2$ (其中 1 在 3 前有一个逆序，1 在 2 前也有一个逆序)，等等. 右下角的第一个下标 2, 1, 3 分别表示第一、二、三行的顺序，因之应每项先写成 $a_{2p}a_{1q}a_{3r}$ 后，再来比较列的逆序数. 所以，在考虑各项的逆序数时，要认清(13)中 1, 2, …, n 所表示的具体意义，决不可形式地拘泥于 1, 2, …, n 这 n 个数字的大小.

练 习 一

1. 已知 $abc\neq 0$, 解方程组:

$$\begin{cases} bx-ay=-2ab, \\ -2cy+3bz=bc, \\ cx+az=0. \end{cases}$$

2. n 级行列式(13)中的项 $a_{1n}a_{2,n-1}a_{3,n-2}\cdots a_{n-1,2}a_{n1}$ 应带什么正负号?

3. 证明: 在一个 n 级行列式中，如果等于零的元素的个数大于 n^2-n，那末这行列式等于零.

第 2 节 行列式的基本性质

在 $n=2$、3 时，可以直接由定义去计算行列式的值，但在

n 较大时，例如 $n=10$，直接由定义去计算 10 级行列式的值，就得计算 $10!=3628800$ 个项，其计算的工作量将非常繁重. 于是，实际要计算行列式时，必须采取另外的途径. 为此，必须先研究行列式有哪些性质，以便利用行列式之性质，来实际地去计算行列式的值.

前已指出：n 级行列式是指有 $n!$ 个项的代数和，每一项含且仅含每行每列的一个元；至于各项的符号是这样规定的，即将该项的 n 个因子(元素)按行的自然顺序排写(也就是把第一、二、…、n 行的元分别写为该项的第一、二、…、n 个因子)，然后再来看这 n 个元素所在列的逆序数的奇或偶而冠以负号"$-$"或正号"$+$". 于是，自然会提出这样一个问题，即从 n 级行列式之每行每列各取一个元之项的 n 个元素若按列的自然顺序排写后，再来看它们所在行的逆序数的奇偶性而冠以负号或正号，所得结果是一致的吗?

可以猜测，上面的问题似可有肯定的答案，也就是要证明行列式(13)与下面的行列式

$$\begin{vmatrix} a_{11} & a_{21} & \cdots & a_{n1} \\ a_{12} & a_{22} & \cdots & a_{n2} \\ \vdots & \vdots & & \vdots \\ a_{1n} & a_{2n} & \cdots & a_{nn} \end{vmatrix} = \sum_{(q_1,q_2,\cdots,q_n)} (-1)^{[q_1,q_2,\cdots,q_n]} a_{q_11}a_{q_22}\cdots a_{q_nn} \qquad (14)$$

相等. 注意行列式(14)是从行列式(13)的行变成列，而将列变成行，但不改变行或列中各元素的排列之顺序而成的. 我们把(14)称为(13)的**转置行列式**. 当然，(13)也是(14)的转置行列式. 对于行列式 D，我们把它的转置行列式记成 D'.

要证明一行列式与它的转置行列式相等，自然会牵涉到排列的次序变更时，它的逆序数应该发生怎样的变化. 对此，

我们有下面的引理:

引理 使某排列中任意两个数字互换，则这排列之逆序数增加或减少一个奇数. 也就是说，互换排列中任两个数字时，排列的奇偶性要变更.

证明 (一)先讨论一种特别情形，即互换的两个数字是排列中相邻的两个数字.

假定原来的排列是

$$\cdots, i, j, \cdots,$$

其中的点子"…"都表示数. 现在只把 i 与 j 互换，得到一个新的排列

$$\cdots, j, i, \cdots;$$

在这两个排列中，数字 i 与 j 和其余的数字所成的逆序显然没有变动. 如果 i 相对 j 在原排列中没有逆序，则它们在新排列中就有一个逆序，即增加了一个逆序数. 如果 i 相对 j 在原排列中已有逆序，则在新排列中就失掉了这逆序，即减少了一个逆序数. 这就是说，当交换两个相邻的数字时，引理是成立的.

(二)再考虑一般情形，即假定互相交换的两数字 i 与 j 中间还有 s 个数字($s>0$)，即排列形状是

$$\cdots, i, k_1, k_2, \cdots, k_s, j, \cdots \tag{15}$$

时，i 与 j 交换后所成的新排列为

$$\cdots, j, k_1, k_2, \cdots, k_s, i, \cdots. \tag{16}$$

这时，可以用下面的方法从排列(15)变到排列(16).

首先，交换排列(15)中两个相邻数字 i 与 k_1 而得一个新排列，再在这新排列中互换 i 与 k_2 这两个相邻的数字又得一新排列，继续这样做下去，直到把 i 与 k_s 互换，再又把 i 与 j 互换，就得到

$$\cdots, k_1, k_2, \cdots, k_s, j, i, \cdots, \tag{17}$$
其中的点子以及点子的次序都和(15)中的完全一样. 最后,再从排列(17)中把 j 顺次与它前面的 $k_s, \cdots, k_2, k_1$ 一个挨一个地互换,就得到(16). 从而可知,由(15)到(16),是从(15)开始通过互相交换相邻两数字共 $s+1$ 回而得(17),然后再从(17)起通过互相交换相邻两数字共 s 回就得到了(16),总起来讲,就是互相交换两个相邻数字共 $2s+1$ 回,可从(15)得到(16). 因为交换两个相邻数字,排列的奇偶性变更了,因此奇偶性变更了奇数 $2s+1$ 回,结果就变更了原排列的奇偶性.

至此,引理完全得证.

有了这引理,就不难证明上面我们提出的问题,即:

定理 1 *行列式和它的转置行列式相等.*

证明 若能证明行列式(13)的任意项都在(14)中出现,那末可用同样方法证明(14)之任意项也出现在(13)中,因此(13)就等于(14),问题就解决了.

取行列式(13)之任一项

$$(-1)^{[p_1, p_2, \cdots, p_n]} a_{1p_1} a_{2p_2} \cdots a_{np_n}.$$

将这个项的 n 个因子 $a_{1p_1}, a_{2p_2}, \cdots, a_{np_n}$ 的次序重新排列一下,使它们的右下角的第二个指标还原成自然排列 $1, 2, \cdots, n$,这时它们的右下角的第一个指标当然就变成了 $1, 2, \cdots, n$ 的某一个排列 $q_1, q_2, \cdots, q_n$. 于是,就有

$$(-1)^{[p_1, p_2, \cdots, p_n]} a_{1p_1} a_{2p_2} \cdots a_{np_n} = (-1)^{[p_1, p_2, \cdots, p_n]} a_{q_1 1} a_{q_2 2} \cdots a_{q_n n}. \tag{18}$$

若能证明

$$(-1)^{[p_1, p_2, \cdots, p_n]} = (-1)^{[q_1, q_2, \cdots, q_n]}, \tag{19}$$

则 $(-1)^{[p_1, p_2, \cdots, p_n]} a_{1p_1} a_{2p_2} \cdots a_{np_n}$ 即为(14)之一项,问题就解决了. 但要证明(19),就只需证明 $[p_1, p_2, \cdots, p_n]$ 与 $[q_1, q_2, \cdots,$

$q_n]$或同为奇数或同为偶数就可以了.

我们知道:从(18)的左边到右边的过程,是可以从交换若干回两个相邻的因子来完成. 但每一次这样的互换, 显然会使第一个下标所成之排列中的两个相邻数字互换了, 同时也使第二个下标所成之排列中的两个相邻数字互换了. 当(18)的左边过渡到右边时, 第一个下标是由排列 1, 2, …, n 变成了 $q_1, q_2, \cdots, q_n$; 而第二个下标则由排列 $p_1, p_2, \cdots, p_n$ 变成了自然排列 1, 2, …, n. 假如从(18)的左边过渡到右边所需要的互换两个相邻因子的回数为 λ, 因为从 1, 2, …, n 变成 $q_1, q_2, \cdots, q_n$ 与从 $p_1, p_2, \cdots, p_n$ 变成 1, 2, …, n 所需要的互换两个相邻数字的回数都可以取成 λ 回, 所以由引理, 就得到

$$
\begin{aligned}
[p_1, p_2, \cdots, p_n] &= [1, 2, \cdots, n] + s_1 + s_2 + \cdots + s_\lambda \\
&= s_1 + s_2 + \cdots + s_\lambda, \\
[q_1, q_2, \cdots, q_n] &= [1, 2, \cdots, n] + t_1 + t_2 + \cdots + t_\lambda \\
&= t_1 + t_2 + \cdots + t_\lambda,
\end{aligned}
$$

式中 $s_1, s_2, \cdots, s_\lambda, t_1, t_2, \cdots, t_\lambda$ 都是奇数(正或负, 实际上都等于 $+1$ 或 -1). 于是, 当 λ 为奇数时, $[p_1, p_2, \cdots, p_n]$ 与 $[q_1, q_2, \cdots, q_n]$ 同是奇数; 当 λ 为偶数时, $[p_1, p_2, \cdots, p_n]$ 与 $[q_1, q_2, \cdots, q_n]$ 则同为偶数. 因之, $[p_1, p_2, \cdots, p_n]$ 与 $[q_1, q_2, \cdots, q_n]$ 的奇偶性相同, 也就说明了行列式(13)之任一项也是(14)的一项. 定理 1 完全得证.

有了定理 1, 即可断言行列式之行与列有同等的功效, 也就是说, 当行列式 D 具有关于行的性质 ε 时, 则 D 也必相应地具有关于列的性质 ε(这是因为, 令 D' 表示 D 的转置行列式, 由假设行列式具行性质 ε, 故 D' 具行性质 ε, 但 D' 的行就是 D 的列, 因而 D 具列性质); 反之亦然.

所以，今后研究行列式的性质时，我们就只着重研究行.

我们知道：行列式的定义中，一个关键问题就是排列的奇偶性. 引理又告诉我们：互换一排列中任两个数字时，这排列的奇偶性就要改变了. 于是自然会问：一行列式之任两个行互相掉换一下，这行列式的值怎样改变呢？容易猜想，其结果只是改变符号的正负，也就是有：

定理 2 将行列式的任两个行(列)互换，行列式之值只变符号.

证明 将行列式(11)用 D 来表示. 互换它的第 i, j 两行($i<j$)，并不动其它的各行，就得一新行列式

$$\Delta=\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{j1} & a_{j2} & \cdots & a_{jn} \\ \vdots & \vdots & & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}\begin{matrix} \\ \\ (i) \\ \\ (j) \\ \\ \\ \end{matrix}$$

(旁边括号内的数 i 与 j 表示行的序数). 因为 Δ 中列的标记与在 D 中的标记一样，即第一、二、…、n 列仍分别用原来的第二个右下标 1, 2, …, n 来标记；但行的标记就不同了，即第一个下标 j, i 分别表示 Δ 的第 i 行与第 j 行，并且由于

$$\begin{aligned}&[p_1, \cdots, p_i, \cdots, p_j, \cdots, p_n]\\ &\quad=[p_1, \cdots, p_j, \cdots, p_i, \cdots, p_n]+\text{奇数},\end{aligned}$$

所以行列式 Δ 的任一项

$$\begin{aligned}&(-1)^{[p_1,\cdots,p_i,\cdots,p_j,\cdots,p_n]}a_{1p_1}\cdots a_{jp_i}\cdots a_{ip_j}\cdots a_{np_n}\\ &\quad=(-1)^{[p_1,\cdots,p_i,\cdots,p_j,\cdots,p_n]}a_{1p_1}\cdots a_{ip_j}\cdots a_{jp_i}\cdots a_{np_n}\\ &\quad=(-1)\{(-1)^{[p_1,\cdots,p_j,\cdots,p_i,\cdots,p_n]}a_{1p_1}\cdots a_{ip_j}\cdots a_{jp_i}\cdots a_{np_n}\},\end{aligned}$$

括号{ }内的式于是行列式 D 中的一项. 此即表明，Δ 之任

一项是 D 之一项的反号. 这就证得了 $\Delta=-D$, 定理 2 证完.

由定理 2, 马上得知下面的

推论 有两行(列)完全相同的行列式等于零.

证明 若行列式 D 的第 i, j 两行(列)完全一样, 则一方面互换第 i, j 两行(列)时, 显然 D 没有发生变化; 可另一方面, 据定理 2, 它又变为 $-D$, 故有 $D=-D$, 即有 $2D=0$, $D=0$. 证毕.

又, 由行列式的定义, 还可直接地得知下列性质:

性质 1 若行列式有一行(列)完全是零, 那末这行列式就等于零.

证明 例如第 i 行(列)的元素完全为零, 则因行列式之各项都有第 i 行(列)的某元素为因子, 故这时行列式之各项都等于零, 因之这行列式自身也等于零.

性质 2 将行列式的某一行(列)中所有元素都同用数 k 去乘, 其结果就等于用数 k 去乘这个行列式.

证明 若用 k 乘行列式 D 之第 i 行(列)的所有元素, 因为 D 之每一项都含有一个且仅一个第 i 行(列)的元素, 所以用 k 乘 D 之第 i 行(列)的所有元素就无异乎是把 D 的每一项都增加了一个因子 k, 于是其结果就等于原行列式 D 乘以数 k.

由性质 2, 即可得到:

推论 1 行列式中某一行(列)的所有元素的公因子, 可以提到行列式符号的外面.

再借助定理 2, 又得到:

推论 2 行列式中若有两个行(列)成比例, 则这行列式等于零.

证明 若行列式 D 之第 j 行(列)的元素为第 i 行(列)的

对应元素的 k 倍 $(i\neq j)$，则当提出第 j 行(列)的公因子 k 到行列式符号的外面时，就得到一个有两行(列)完全相同的行列式，因而据定理 2 的推论，知 $D=0$.

由行列式之定义，可知属于同一行的两个元素决不会出现在同一个项内，故若把行列式看做任何一行之元素的函数(例如第 i 行之元素 $a_{i1}, a_{i2}, \cdots, a_{in}$ 的函数)，则它为 $a_{i1}, a_{i2}, \cdots, a_{in}$ 的一次齐次式．换句话说，可写做

$$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} = a_{i1}A_{i1}+a_{i2}A_{i2}+\cdots+a_{in}A_{in}$$

的形状，其中 $A_{i1}, A_{i2}, \cdots, A_{in}$ 不含第 i 行的元素．对列说来，也是这样．据此，即可证明下面的：

性质 3 如果行列式中第 i 行(列)的所有元素都可以写成两项的和的形状：

$$a_{ij}(\text{或 } a_{ji})=b_j+c_j, \quad (j=1, 2, \cdots, n)$$

那末，这行列式等于两个行列式的和，在这两个行列式的第 i 行(列)的元素一个是由 b_j 所组成，另一个是由 c_j 所组成，其他各行(列)的元素都同原来行列式中的一样，即

$$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots \\ b_1+c_1 & b_2+c_2 & \cdots & b_n+c_n \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots \\ b_1 & b_2 & \cdots & b_n \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} + \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots \\ c_1 & c_2 & \cdots & c_n \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}.$$

证明 等号左端的行列式为

$$(b_1+c_1)A_{i1}+(b_2+c_2)A_{i2}+\cdots+(b_n+c_n)A_{in},$$

它可分解为

$$(b_1A_{i1}+b_2A_{i2}+\cdots+b_nA_{in})+(c_1A_{i1}+c_2A_{i2}+\cdots+c_nA_{in}),$$

恰好等于上述等号右端的两个行列式之和，证毕．

由性质 3，并利用上推论 2，又可得：

性质 4 将行列式之某行(列)的 k 倍加到另一行(列)上，所得的新行列式与原行列式相等.

证明 设将行列式 D 之第 i 行的 k 倍加到第 j 行上 $(j\neq i)$ 所成之新行列式记为 D_1，即

$$D_1=\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & \vdots & & \vdots \\ a_{j1}+ka_{i1} & a_{j2}+ka_{i2} & \cdots & a_{jn}+ka_{in} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix},$$

于是，据性质 3，并由性质 2 的推论 2，就知道：

$$D_1=\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & \vdots & & \vdots \\ a_{j1} & a_{j2} & \cdots & a_{jn} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}+\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & \vdots & & \vdots \\ ka_{i1} & ka_{i2} & \cdots & ka_{in} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}$$

$$=D+0=D,$$

证毕.

[例 1] 设行列式

$$D=\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix},$$

试问行列式

$$
\Delta=\begin{vmatrix}
a_{1p_1} & a_{1p_2} & \cdots & a_{1p_n}\\
a_{2p_1} & a_{2p_2} & \cdots & a_{2p_n}\\
\vdots & \vdots & & \vdots\\
a_{np_1} & a_{np_2} & \cdots & a_{np_n}
\end{vmatrix}
$$

或
$$
\Delta=\begin{vmatrix}
a_{p_11} & a_{p_12} & \cdots & a_{p_1n}\\
a_{p_21} & a_{p_22} & \cdots & a_{p_2n}\\
\vdots & \vdots & & \vdots\\
a_{p_n1} & a_{p_n2} & \cdots & a_{p_nn}
\end{vmatrix}
$$

与 D 有怎样的关系？这里 $p_1, p_2, \cdots, p_n$ 是 $1, 2, \cdots, n$ 的某个排列.

解 设排列 $p_1, p_2, \cdots, p_n$ 是从 $1, 2, \cdots, n$ 中互换两个数字共 k 回而成，由定理 2，可知 $\Delta=(-1)^kD$. 由引理又知道

$$
\begin{aligned}
[p_1, p_2, \cdots, p_n]&=[1, 2, \cdots, n]+s_1+s_2+\cdots+s_k\\
&=s_1+s_2+\cdots+s_k,
\end{aligned}
$$

式中每个 s_i 为奇数，故 $[p_1, p_2, \cdots, p_n]$ 与 k 之差为偶数，因而

$$
(-1)^k=(-1)^{[p_1,p_2,\cdots,p_n]},
$$

$$
\Delta=(-1)^{[p_1,p_2,\cdots,p_n]}\cdot D.
$$

同理，
$$
\Delta=(-1)^{[p_1,p_2,\cdots,p_n]}\cdot D.
$$

这就是所要求的关系.

[例 2] 证明：

$$
D=\begin{vmatrix}
a_{11} & a_{12} & \cdots & a_{1n}\\
a_{21} & a_{22} & \cdots & a_{2n}\\
\vdots & \vdots & & \vdots\\
a_{n1} & a_{n2} & \cdots & a_{nn}
\end{vmatrix}
=\sum(-1)^{[i_1,i_2,\cdots,i_n]+[j_1,j_2,\cdots,j_n]}\cdot a_{i_1j_1}a_{i_2j_2}\cdots a_{i_nj_n},
$$

式中 $i_1, i_2, \cdots, i_n$ 与 $j_1, j_2, \cdots, j_n$ 都是 $1, 2, \cdots, n$ 的排列，但 $\sum$ 表示取这样一些排列之积的和，即 $i_1, i_2, \cdots, i_n$ 为 $1, 2, \cdots, n$ 的某一个固定的排列，而 $j_1, j_2, \cdots, j_n$ 得取

1, 2, …, n 的所有的排列(共 $n!$ 个),或者是 $j_1, j_2, \cdots, j_n$ 为 1, 2, …, n 的某一个固定的排列, 而 $i_1, i_2, \cdots, i_n$ 得取 1, 2, …, n 的所有的排列(共 $n!$ 个).

证明 不失一般性,可只考虑 $i_1, i_2, \cdots, i_n$ 为 1, 2, …, n 的某固定排列, 而 $j_1, j_2, \cdots, j_n$ 得取 1, 2, …, n 之共 $n!$ 个排列.

设将 $D=\sum\limits_{(p_1,\cdots,p_n)}(-1)^{[p_1,p_2,\cdots,p_n]}\cdot a_{1p_1}a_{2p_2}\cdots a_{np_n}$ 之任一项 $(-1)^{[p_1,p_2,\cdots,p_n]}a_{1p_1}a_{2p_2}\cdots a_{np_n}$ 的 n 个(因子)元素 $a_{1p_1}, a_{2p_2}, \cdots, a_{np_n}$ 重新排列一下, 使为第 $i_1, i_2, \cdots, i_n$ 行的元分别放在积因子的第一、二、…、n 个的位置上,于是,

$$(-1)^{[p_1,p_2,\cdots,p_n]}a_{1p_1}a_{2p_2}\cdots a_{np_n}=(-1)^{[p_1,p_2,\cdots,p_n]}a_{i_1j_1}a_{i_2j_2}\cdots a_{i_nj_n}$$

形, 即原来的下标 1, 2, …, n 变成了 $i_1, i_2, \cdots, i_n$, 而 $p_1, p_2, \cdots, p_n$ 就变成了 $j_1, j_2, \cdots, j_n$. 今令由 $a_{1p_1}a_{2p_2}\cdots a_{np_n}$ 变为 $a_{i_1j_1}a_{i_2j_2}\cdots a_{i_nj_n}$, 可假定是从 $a_{1p_1}, a_{2p_2}, \cdots, a_{np_n}$ 中互换二个因子共 k 次而成, 于是下标 1, 2, …, n 变成 $i_1, i_2, \cdots, i_n$ 以及 $p_1, p_2, \cdots, p_n$ 变成 $j_1, j_2, \cdots, j_n$ 都可以看做是互换两数字共 k 次而成,故

$$[i_1, i_2, \cdots, i_n]=[1, 2, \cdots, n]+t_1+t_2+\cdots+t_k,$$

$$[j_1, j_2, \cdots, j_n]=[p_1, p_2, \cdots, p_n]+s_1+s_2+\cdots+s_k,$$

每 t_i 与 s_i 都是奇数,因而

$$\begin{aligned}&[i_1, i_2, \cdots, i_n]+[j_1, j_2, \cdots, j_n]\\&=[1, 2, \cdots, n]+[p_1, p_2, \cdots, p_n]+\sum_{i=1}^{k}(s_i+t_i)\\&=[p_1, p_2, \cdots, p_n]+\text{偶数},\end{aligned}$$

说明了 $(-1)^{[p_1,p_2,\cdots,p_n]}=(-1)^{[i_1,i_2,\cdots,i_n]+[j_1,j_2,\cdots,j_n]}$,

故 $$\begin{aligned}D&=\sum(-1)^{[p_1,p_2,\cdots,p_n]}\cdot a_{1p_1}a_{2p_2}\cdots a_{np_n}\\&=\sum(-1)^{[i_1,i_2,\cdots,i_n]+[j_1,j_2,\cdots,j_n]}a_{i_1j_1}a_{i_2j_2}\cdots a_{i_nj_n}.\end{aligned}$$

即例 2 获证.

［例 3］ 设

$$D=\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}$$

中每一个元 a_{ij} 都是 x 的可微函数，证明:

$$\frac{dD}{dx}=\begin{vmatrix} \frac{d}{dx}a_{11} & \frac{d}{dx}a_{12} & \cdots & \frac{d}{dx}a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}+\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \frac{d}{dx}a_{21} & \frac{d}{dx}a_{22} & \cdots & \frac{d}{dx}a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}+\cdots$$

$$+\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ \frac{d}{dx}a_{n1} & \frac{d}{dx}a_{n2} & \cdots & \frac{d}{dx}a_{nn} \end{vmatrix}.$$

证明 由 $D=\sum(-1)^{[p_1,p_2,\cdots,p_n]}a_{1p_1}a_{2p_2}\cdots a_{np_n}$ 得

$$\frac{d}{dx}D=\sum(-1)^{[p_1,p_2,\cdots,p_n]}\left(\frac{d}{dx}a_{1p_1}\right)a_{2p_2}\cdots a_{np_n}$$

$$+\sum(-1)^{[p_1,p_2,\cdots,p_n]}a_{1p_1}\left(\frac{d}{dx}a_{2p_2}\right)\cdots a_{np_n}+\cdots$$

$$+\sum(-1)^{[p_1,p_2,\cdots,p_n]}a_{1p_1}a_{2p_2}\cdots\left(\frac{d}{dx}a_{np_n}\right)$$

$$=\begin{vmatrix} \frac{d}{dx}a_{11} & \cdots & \frac{d}{dx}a_{1n} \\ a_{21} & \cdots & a_{2n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix}+\begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \frac{d}{dx}a_{21} & \cdots & \frac{d}{dx}a_{2n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix}+\cdots$$

$$+\begin{vmatrix} a_{11} & \cdots & a_{1n} \\ a_{21} & \cdots & a_{2n} \\ \vdots & & \vdots \\ \frac{d}{dx}a_{n1} & \cdots & \frac{d}{dx}a_{nn} \end{vmatrix}.$$

如果将 D 之第 i 行各元求导而当做一新行列式的第 i 行，其余各行之元仍保持 D 中的一样，这样所得之新行列式记以 $D_i(i=1, 2, \cdots, n)$，由例 3 知：

$$\frac{d}{dx}D=D_1+D_2+\cdots+D_n.$$

这给出了求一个 n 级行列式 D 的导数的方法．当然，上述结论中，也可以将“行”字改为“列”字，得求导之另一方法．

练 习 二

1．已知 204，527，255 三数都能用 17 除尽，试证三级行列式

$$\begin{vmatrix} 2 & 0 & 4 \\ 5 & 2 & 7 \\ 2 & 5 & 5 \end{vmatrix}$$

也能用 17 除尽．

2．计算

$$\begin{vmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 5 \\ 3 & 4 & 5 & 6 \\ 4 & 5 & 6 & 7 \end{vmatrix},$$

并推广到一般，计算 n 级行列式

$$\begin{vmatrix} 1 & 2 & 3 & \cdots & n-1 & n \\ 2 & 3 & 4 & \cdots & n & n+1 \\ 3 & 4 & 5 & \cdots & n+1 & n+2 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ n-1 & n & n+1 & \cdots & 2n-3 & 2n-2 \\ n & n+1 & n+2 & \cdots & 2n-2 & 2n-1 \end{vmatrix}.$$

3. 设 $n\geqslant 2$，求

$$D_n=\begin{vmatrix} 1+x_1y_1 & 2+x_1y_2 & \cdots & n+x_1y_n \\ 1+x_2y_1 & 2+x_2y_2 & \cdots & n+x_2y_n \\ \vdots & \vdots & & \vdots \\ 1+x_ny_1 & 2+x_ny_2 & \cdots & n+x_ny_n \end{vmatrix}.$$

4. 证明：n 级行列式的 $n!$ 个项中，有一半的项$\left(即\ \frac{1}{2}n!\ 个项\right)$是取"+"号，另一半的项是取"－"号.

5. n 级行列式

$$D_n=\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}$$

中，如果 $a_{ij}=a_{ji}$，则叫做对称的；当 $a_{ij}=-a_{ji}$ 时叫做反对称的. 试证明：奇数级反对称行列式等于零.

第 3 节　行列式的计算

当 n 相当大时，由定义直接去求 n 级行列式的值是一件非常麻烦且困难的工作，那么，怎样实际地去计算一个行列式呢？

对于二级、三级行列式，在第 1 节中早已说过，它们可以用对角线的方法来计算. 可是，这样的法则不能推广到级较高的行列式. 现在，用前述行列式的性质来降低行列式之级，从而简化行列式的计算过程，这就需要研究行列式中所包含的级较低的一些行列式.

一般，在 n 级行列式

$$D=\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}$$

中，删去元素 a_{ij} 所在的第 i 行和第 j 列后而得到的一个 $n-1$ 级行列式，叫做元素 a_{ij} 的余子式，简称为 a_{ij} 的子式，通常表以 M_{ij}，并把

$$A_{ij}=(-1)^{i+j}\cdot M_{ij}$$

定义为元 a_{ij} 的代数余子式．显然，M_{ij} 为包含在 D 内的一个级较低的行列式（$n-1$ 级），因之又称为 D 内的一个 $n-1$ 级子式．

上面所谓降低行列式之级而简化行列式的计算过程，指的是下面的：

定理 1　n 级行列式

$$D=\begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix}$$

等于它的任一行（列）中所有元素与它们对应的代数余子式的乘积之和，即

$$D=a_{i1}A_{i1}+a_{i2}A_{i2}+\cdots+a_{in}A_{in}, \tag{20}$$

或

$$D=a_{1i}A_{1i}+a_{2i}A_{2i}+\cdots+a_{ni}A_{ni}. \tag{21}$$

（其中 $i=1, 2, \cdots, n$）．

定理 1 的重要意义是把 n 级行列式的计算简化为 $n-1$ 级行列式的计算．于是，再用一次定理 1 又可降为 $n-2$ 级，继续这样降低级数，直到容易计算为止．

为证明定理 1，先来证明下面的：

引理　行列式 D 中任一元素与它的代数余子式的乘积，是行列式 D 中一部分的项．

证明　先证明当元素在 D 之左上角时的特殊情形，即要证明 $a_{11}A_{11}\subset D$．

由于 $A_{11}=M_{11}$，故

$$a_{11}A_{11}=a_{11}M_{11}=a_{11}\begin{vmatrix} a_{22} & \cdots & a_{2n} \\ \vdots & & \vdots \\ a_{n2} & \cdots & a_{nn} \end{vmatrix}$$

$$=\sum_{(p_2,\cdots,p_n)}(-1)^{[p_2,\cdots,p_n]}\cdot a_{11}a_{2p_2}a_{3p_3}\cdots a_{np_n},$$

式中$(p_2, p_3, \cdots, p_n)$为$(2, 3, \cdots, n)$的一排列，而$[p_2, p_3, \cdots, p_n]$是排列$p_2, p_3, \cdots, p_n$依2, 3, …, n的顺序比较所得的逆序数. 于是1, $p_2, p_3, \cdots, p_n$为1, 2, 3, …, n之一排列，且显有$[1, p_2, p_3, \cdots, p_n]=[p_2, p_3, \cdots, p_n]$，故上式右端之$\Sigma$即为$D$中一部分的项，亦即$a_{11}A_{11}\subset D$，此即表明，引理在这特殊情况时是成立的.

再证明一般情形，即要证$a_{ij}A_{ij}\subset D$. 事实上，先将第i行与它上面的第$i-1$行、第$i-2$行、…、第二行、第一行挨次地互换，共需要互换$i-1$回，可将原第i行变为新行列式的第一行；然后，再将这新行列式之第j列与它前面的第$j-1$列、第$j-2$列、…、第一列又挨次地互换，共需互换$j-1$回，最后得到了另一个新行列式，记为Δ. 显然，a_{ij}在Δ的左上角，而a_{ij}在Δ内的代数余子式就是a_{ij}在D内的余子式M_{ij}. 于是，根据已证过了的特殊情况，就应有$a_{ij}M_{ij}\subset\Delta$，据此即得$a_{ij}\cdot(-1)^{i+j}M_{ij}\subset(-1)^{i+j}\Delta$，即$a_{ij}A_{ij}\subset D$.

引理证完.

运用引理的结论，不难证明定理1，即(20)、(21)式.

定理1的证明 从n个积$a_{i1}A_{i1}, a_{i2}A_{i2}, \cdots, a_{in}A_{in}$中任取两个$a_{is}A_{is}$与$a_{it}A_{it}$. 根据引理，$a_{is}A_{is}$与$a_{it}A_{it}$分别都包含了$D$之$(n-1)!$个项. 而包含在$a_{is}A_{is}$中的$D$之$(n-1)!$个项里每一项都含有$D$之第$i$行的元素$a_{is}$，包含在$a_{it}A_{it}$中的$D$之$(n-1)!$个项里每一项都含有$D$之第$i$行的元素$a_{it}$，又$D$的每一项必包含其第$i$行的唯一个元，故含元$a_{is}$为因子的

项决不会再含元 a_{it} 为因子，同理含元 a_{it} 为因子的项也不含元 a_{is} 为因子，于是 $a_{is}A_{is}$ 与 $a_{it}A_{it}$ 没有共同的项，这就说明了 n 个积 $a_{i1}A_{i1}$，$a_{i2}A_{i2}$，…，$a_{in}A_{in}$ 中任两个都没有共同的项．故 $a_{i1}A_{i1}, a_{i2}A_{i2}, \cdots, a_{in}A_{in}$ 合并起来共包含了 D 的 $n\cdot(n-1)!=n!$ 个项，即包含了 D 之全部的项，也就是说

$$D=a_{i1}A_{i1}+a_{i2}A_{i2}+\cdots+a_{in}A_{in}.$$

同理，$D=a_{1t}A_{1t}+a_{2t}A_{2t}+\cdots+a_{nt}A_{nt}$．定理 1 证完．

如用简洁的语言表达定理 1，它的意义就是说：*一个行列式可按照它的任一行展开，也可按照它的任一列展开．*

在(20)中，如将 a_{i1}，a_{i2}，…，a_{in} 分别换为 a_{j1}，a_{j2}，…，$a_{jn}(i\neq j)$，但 A_{i1}，A_{i2}，…，A_{in} 不改换，而得到的式子 $a_{j1}A_{i1}+a_{j2}A_{i2}+\cdots+a_{jn}A_{in}$，显然是等于另一个新行列式，这新行列式只将 D 之第 i 行改换为 a_{j1}，a_{j2}，…，a_{jn}，也就是说新行列式的第 i，j 两行相同(都为 a_{j1}，a_{j2}，…，a_{jn})，故必等于零，即 $a_{j1}A_{i1}+a_{j2}A_{i2}+\cdots+a_{jn}A_{in}=0(i\neq j)$．同样，在(21)中，也可得到 $a_{1j}A_{1i}+a_{2j}A_{2i}+\cdots+a_{nj}A_{ni}=0(j\neq i)$．因而，结合定理 1 的结论，可有：

定理 2 *行列式 D 可依其任一行(或列)展开，亦即 D 之任一行(或列)与其对应之代数余子式之乘积的和就等于 D 自身．但是，D 之任一行(或列)与另外一行(或列)之对应的代数余子式之乘积的和恒等于零．*

定理 2 用式子来表示，就是

$$\sum_{k=1}^{n} a_{ik}A_{jk}=\begin{cases}D, & \text{当 } j=i \text{ 时},\\ 0, & \text{当 } j\neq i \text{ 时};\end{cases}$$

$$\sum_{k=1}^{n} a_{ki}A_{kj}=\begin{cases}D, & \text{当 } j=i \text{ 时},\\ 0, & \text{当 } j\neq i \text{ 时}.\end{cases}$$

定理 1 告诉我们怎样用降级的方法来计算行列式，但在

实际运用中仍然是很麻烦的，因为需要计算许多降了级的行列式．所以，在运用这方法时，还要结合前节所讲的行列式之性质，使某行(列)之元素尽可能多地变成零，便于按这行(列)展开时不必去计算元素 0 所在之行与列的代数余子式．

［例 1］ 证明范得蒙(Vandermonde)行列式：

$$D_n=\begin{vmatrix} 1 & 1 & 1 & \cdots & 1 \\ x_1 & x_2 & x_3 & \cdots & x_n \\ x_1^2 & x_2^2 & x_3^2 & \cdots & x_n^2 \\ \vdots & \vdots & \vdots & & \vdots \\ x_1^{n-1} & x_2^{n-1} & x_3^{n-1} & \cdots & x_n^{n-1} \end{vmatrix}=\prod_{1\leq j<i\leq n}(x_i-x_j).$$

证明 用数学归纳法．因为

$$D_2=\begin{vmatrix} 1 & 1 \\ x_1 & x_2 \end{vmatrix}=x_2-x_1=\prod_{1\leq j<i\leq 2}(x_i-x_j),$$

所以 $n=2$ 时上式成立．

现在假定对于 $n-1$ 级范得蒙行列式上式成立，在这假设下，我们来证明对于 n 级范得蒙行列式上式也成立．

我们设法把 D_n 化为 $n-1$ 级行列式．首先从最后第 n 行减去第 $n-1$ 行的 x_1 倍，再从第 $n-1$ 行减去第 $n-2$ 行的 x_1 倍，继续这样进行，直到第 2 行减去第 1 行的 x_1 倍，我们就得出

$$D_n=\begin{vmatrix} 1 & 1 & 1 & \cdots & 1 \\ 0 & x_2-x_1 & x_3-x_1 & \cdots & x_n-x_1 \\ 0 & x_2(x_2-x_1) & x_3(x_3-x_1) & \cdots & x_n(x_n-x_1) \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & x_2^{n-2}(x_2-x_1) & x_3^{n-2}(x_3-x_1) & \cdots & x_n^{n-2}(x_n-x_1) \end{vmatrix};$$

再按第一列展开，并把每列的公因子提到外面，就有

$$D_n=(x_2-x_1)(x_3-x_1)\cdots(x_n-x_1)\begin{vmatrix}1 & 1 & \cdots & 1\\ x_2 & x_3 & \cdots & x_n\\ \vdots & \vdots & & \vdots\\ x_2^{n-2} & x_3^{n-2} & \cdots & x_n^{n-2}\end{vmatrix};$$

上式右边的行列式是 $n-1$ 级范得蒙行列式，故由归纳法的假设，它应该等于所有 x_i-x_j 的乘积，其中 $2\leqslant j<i\leqslant n$. 因此，

$$D_n=(x_2-x_1)(x_3-x_1)\cdots(x_n-x_1)\cdot\prod_{2\leqslant j<i\leqslant n}(x_i-x_j)$$
$$=\prod_{1\leqslant j<i\leqslant n}(x_i-x_j).$$

例 1 获证.

[例 2] 试求 n 级循环行列式

$$D=\begin{vmatrix}x_0 & x_1 & x_2 & \cdots & x_{n-2} & x_{n-1}\\ x_{n-1} & x_0 & x_1 & \cdots & x_{n-3} & x_{n-2}\\ x_{n-2} & x_{n-1} & x_0 & \cdots & x_{n-4} & x_{n-3}\\ \vdots & \vdots & \vdots & & \vdots & \vdots\\ x_2 & x_3 & x_4 & \cdots & x_0 & x_1\\ x_1 & x_2 & x_3 & \cdots & x_{n-1} & x_0\end{vmatrix}.$$

解 设 $\omega_1(=1)$, ω_2, $\cdots$, ω_n 为全部 n 次单位根（即为 $x^n-1=0$ 之 n 个根）. 今以 ω_i, ω_i^2, $\cdots$, ω_i^{n-2}, ω_i^{n-1} 分别乘 D 之第二、三、$\cdots$、$n-1$、n 列后 $(i=1,\ 2,\ \cdots,\ n)$ 统统加到第一列上去，就得到

$$D=(x_0+\omega_i x_1+\omega_i^2x_2+\cdots+\omega_i^{n-2}x_{n-2}$$
$$+\omega_i^{n-1}x_{n-1})\cdot\begin{vmatrix}1 & x_1 & x_2 & \cdots & x_{n-2} & x_{n-1}\\ \omega_i & x_0 & x_1 & \cdots & x_{n-3} & x_{n-2}\\ \omega_i^2 & x_{n-1} & x_0 & \cdots & x_{n-4} & x_{n-3}\\ \vdots & \vdots & \vdots & & \vdots & \vdots\\ \omega_i^{n-2} & x_3 & x_4 & \cdots & x_0 & x_1\\ \omega_i^{n-1} & x_2 & x_3 & \cdots & x_{n-1} & x_0\end{vmatrix},$$

由于 D 可看作是 n 个变量 $x_0, x_1, \cdots, x_{n-1}$ 的 n 次齐次多项式，上式表明 D 可被 n 个互异的 $x_0, x_1, \cdots, x_{n-1}$ 之一次齐次多项式 $x_0+\omega_i x_1+\omega_i^2 x_2+\cdots+\omega_i^{n-2}x_{n-2}+\omega_i^{n-1}x_{n-1}$ 所整除。上式中令 $i=1, 2, \cdots, n$，再比较 x_0^n 之系数，即可得出

$$D=\prod_{i=1}^{n}(x_0+\omega_i x_1+\omega_i^2 x_2+\cdots+\omega_i^{n-2}x_{n-2}+\omega_i^{n-1}x_{n-1}).$$

练　习　三

1. 利用范得蒙行列式，计算：

(i) $$\begin{vmatrix} a^n & (a+1)^n & \cdots & (a+n)^n \\ a^{n-1} & (a+1)^{n-1} & \cdots & (a+n)^{n-1} \\ \vdots & \vdots & & \vdots \\ a & a+1 & \cdots & a+n \\ 1 & 1 & \cdots & 1 \end{vmatrix},$$

(ii) $$\begin{vmatrix} a_1^n & a_1^{n-1}b_1 & \cdots & a_1 b_1^{n-1} & b_1^n \\ a_2^n & a_2^{n-1}b_2 & \cdots & a_2 b_2^{n-1} & b_2^n \\ \vdots & \vdots & & \vdots & \vdots \\ a_{n+1}^n & a_{n+1}^{n-1}b_{n+1} & \cdots & a_{n+1}b_{n+1}^{n-1} & b_{n+1}^n \end{vmatrix},$$

这里的 $a_i\neq 0$，$i=1, 2, \cdots, n+1$。

2. 用数学归纳法，证明：

(i) $$D_n=\begin{vmatrix} a+b & ab & 0 & \cdots & 0 & 0 \\ 1 & a+b & ab & \cdots & 0 & 0 \\ 0 & 1 & a+b & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & a+b & ab \\ 0 & 0 & 0 & \cdots & 1 & a+b \end{vmatrix}=\frac{a^{n+1}-b^{n+1}}{a-b},$$

(ii) $$D_n=\begin{vmatrix} 2\cos\theta & 1 & 0 & \cdots & 0 & 0 \\ 1 & 2\cos\theta & 1 & \cdots & 0 & 0 \\ 0 & 1 & 2\cos\theta & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 2\cos\theta & 1 \\ 0 & 0 & 0 & \cdots & 1 & 2\cos\theta \end{vmatrix}=\frac{\sin(n+1)\theta}{\sin\theta}.$$

3. 证明：n 级行列式

$$\begin{vmatrix} x & -1 & 0 & \cdots & 0 & 0 \\ 0 & x & -1 & \cdots & 0 & 0 \\ 0 & 0 & x & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & x & -1 \\ a_n & a_{n-1} & a_{n-2} & \cdots & a_2 & x+a_1 \end{vmatrix} = x^n + a_1 x^{n-1} + \cdots + a_{n-1}x + a_n.$$

4. 设 n 级行列式 D_n 之主对角线†上的元素全是 a，而在主对角线的紧上及紧下的元素均分别为 b 及 c，其余元全为零，即

$$D_n = \begin{vmatrix} a & b & 0 & 0 & \cdots & 0 & 0 \\ c & a & b & 0 & \cdots & 0 & 0 \\ 0 & c & a & b & \cdots & 0 & 0 \\ 0 & 0 & c & a & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & a & b \\ 0 & 0 & 0 & 0 & \cdots & c & a \end{vmatrix},$$

试求存在于 D_n, D_{n-1} 及 D_{n-2} 间的关系式．并当 $a=1+bc$ 时计算 D_n．

5. 设 $s=a_1+a_2+\cdots+a_n$, $A_k=s-a_k$（其中 $k=1, \cdots, n$），证明：

$$\begin{vmatrix} x-A_1 & a_2 & a_3 & \cdots & a_n \\ a_1 & x-A_2 & a_3 & \cdots & a_n \\ a_1 & a_2 & x-A_3 & \cdots & a_n \\ \vdots & \vdots & \vdots & & \vdots \\ a_1 & a_2 & a_3 & \cdots & x-A_n \end{vmatrix} = x(x-s)^{n-1}.$$

6. 试求下列各行列式：

(i) $\begin{vmatrix} 1 & 2 & 3 & \cdots & n \\ 1 & 2^2 & 3^2 & \cdots & n^2 \\ 1 & 2^3 & 3^3 & \cdots & n^3 \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & 2^n & 3^n & \cdots & n^n \end{vmatrix}$; (ii) $\begin{vmatrix} a_{11}+x & a_{12}+x & \cdots & a_{1n}+x \\ a_{21}+x & a_{22}+x & \cdots & a_{2n}+x \\ \vdots & \vdots & & \vdots \\ a_{n1}+x & a_{n2}+x & \cdots & a_{nn}+x \end{vmatrix}$;

(iii) $\begin{vmatrix} 1 & n & n & \cdots & n & n \\ n & 2 & n & \cdots & n & n \\ n & n & 3 & \cdots & n & n \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ n & n & n & \cdots & n-1 & n \\ n & n & n & \cdots & n & n \end{vmatrix}$; (iv) $\begin{vmatrix} 1 & 2 & 0 & 0 & \cdots & 0 & 0 \\ 1 & 3 & 2 & 0 & \cdots & 0 & 0 \\ 0 & 1 & 3 & 2 & \cdots & 0 & 0 \\ 0 & 0 & 1 & 3 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 3 & 2 \\ 0 & 0 & 0 & 0 & \cdots & 1 & 3 \end{vmatrix}$;

† 从左上角到右下角的对角线叫主对角线。

(v) $\begin{vmatrix} 1+x^2 & x & & & & \\ x & 1+x^2 & x & & & \\ & x & 1+x^2 & \ddots & & \\ & & \ddots & \ddots & \ddots & \\ & & & \ddots & 1+x^2 & x \\ & & & & x & 1+x^2 \end{vmatrix}$.

第4节 克莱姆定理

n级行列式的概念是2级、3级行列式的推广，而2级、3级行列式是来自解线性方程组的，今问n级行列式能否用来解n个未知量的线性方程组？也就是说，用来解n个未知量的线性方程组时，能否得到与前面类似的表达式呢？这节就是回答这个问题，也是本章开头提出要解决的第四个问题.

设有n个未知量$x_1, x_2, \cdots, x_n$的线性方程组

$$\begin{cases} a_{11}x_1+a_{12}x_2+\cdots+a_{1n}x_n=b_1, \\ a_{21}x_1+a_{22}x_2+\cdots+a_{2n}x_n=b_2, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{n1}x_1+a_{n2}x_2+\cdots+a_{nn}x_n=b_n, \end{cases} \tag{22}$$

并假定(22)之系数所组成的行列式

$$D=\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}\neq 0.$$

若(22)有解，并令$(x_1, x_2, \cdots, x_n)$为它的一个解，于是，以$A_{1i}, A_{2i}, \cdots, A_{ni}$分别乘(22)之第一、二、…、$n$个式子($A_{ji}$表示$a_{ji}$在$D$内的代数余子式)，然后再相加，根据第3节的

定理 2, $x_j(j\neq i)$之系数全为零,而有

$$(a_{1i}A_{1i}+a_{2i}A_{2i}+\cdots+a_{ni}A_{ni})x_i=b_1A_{1i}+b_2A_{2i}+\cdots+b_nA_{ni},$$

即 $Dx_i=D_i$. 但 $D_i=b_1A_{1i}+b_2A_{2i}+\cdots+b_nA_{ni}$ 显为将 D 之第 i 列换为 b_1, b_2, $\cdots$, b_n 后所成之新行列式, 故由 $D\neq 0$ 得知 $x_i=\dfrac{D_i}{D}(i=1, 2, \cdots, n)$, 即 (22) 之解必为 $\left(\dfrac{D_1}{D}, \dfrac{D_2}{D}, \cdots, \dfrac{D_n}{D}\right)$. 这说明了: 方程组 (22) 如有解, 则其解必为 $\left(\dfrac{D_1}{D}, \dfrac{D_2}{D}, \cdots, \dfrac{D_n}{D}\right)$, 故是唯一的.

反之, 若将 D_i 写为 $D_i=b_1A_{1i}+b_2A_{2i}+\cdots+b_nA_{ni}$, 又容易知道:

$$\begin{aligned}
&a_{j1}\frac{D_1}{D}+a_{j2}\frac{D_2}{D}+\cdots+a_{jn}\frac{D_n}{D}\\
&=\frac{1}{D}\cdot\{a_{j1}D_1+a_{j2}D_2+\cdots+a_{jn}D_n\}\\
&=\frac{1}{D}\cdot\{a_{j1}(b_1A_{11}+b_2A_{21}+\cdots+b_nA_{n1})\\
&\quad+a_{j2}(b_1A_{12}+b_2A_{22}+\cdots+b_nA_{n2})+\cdots\\
&\quad+a_{jn}(b_1A_{1n}+b_2A_{2n}+\cdots+b_nA_{nn})\}\\
&=\frac{1}{D}\cdot\{b_1\cdot\delta_{j1}D+\cdots+b_j\cdot\delta_{jj}D+\cdots+b_n\cdot\delta_{jn}D\},
\end{aligned}$$

式中 δ_{jt} 为 Kronecker 符号, 即

$$\delta_{jt}=\begin{cases}1, & \text{当 } j=t \text{ 时},\\ 0, & \text{当 } j\neq t \text{ 时},\end{cases}$$

故上式的最后式等于

$$\frac{1}{D}\cdot b_jD=b_j,$$

这就验证了$\left(\dfrac{D_1}{D}, \dfrac{D_2}{D}, \cdots, \dfrac{D_n}{D}\right)$确为方程组(22)的解.

综上所述，就证得了下面的重要结果：

定理 1（克莱姆定理） 含 n 个未知量的 n 个方程的线性方程组(22)，如果它的系数的行列式 $D\neq0$，那末这方程组一定有解，并且只有一个解．它唯一的解是

$$x_1=\frac{D_1}{D},\ x_2=\frac{D_2}{D},\ \cdots,\ x_n=\frac{D_n}{D},$$

式中 $D_j(j=1,\ 2,\ \cdots,\ n)$ 是将 D 的第 j 列之元素 $a_{1j},\ a_{2j},\ \cdots,\ a_{nj}$ 换成常数项 $b_1,\ b_2,\ \cdots,\ b_n$ 所得到的行列式．

显然，克莱姆定理是前述二个、三个未知量的线性方程组解的表达式的推广，这正是我们所希望的结果．

再讨论方程组(22)的一个特殊情形．方程组(22)右边的常数项 $b_1,\ b_2,\ \cdots,\ b_n$ 如果都是零，即

$$\begin{cases}a_{11}x_1+a_{12}x_2+\cdots+a_{1n}x_n=0,\\ a_{21}x_1+a_{22}x_2+\cdots+a_{2n}x_n=0,\\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\ a_{n1}x_1+a_{n2}x_2+\cdots+a_{nn}x_n=0.\end{cases}\tag{23}$$

这样的线性方程组叫做齐次线性方程组．它作为(22)的特例，有其广泛的应用，因此齐次线性方程组在线性方程组的理论中占有重要的地位．

显然，$x_1=0,\ x_2=0,\ \cdots,\ x_n=0$ 为(23)的解．线性方程组的解如果全都等于零就叫做零解，如果不全等于零就叫做非零解．因此，齐次线性方程组必有零解，非齐次线性方程组的解都为非零解．

假设(23)的行列式 $D\neq0$，因这时 $D_1=D_2=\cdots=D_n=0$，故由克莱姆定理，(23)只有唯一个零解．因之，若(23)具有非零解，必然 $D=0$，也就是证得了：

定理 2 含 n 个未知量的 n 个线性方程的齐次方程组

(23)如有非零解，则它的系数的行列式等于零.

注意：克莱姆定理只有在 $D\neq 0$ 时才能成立. 当 $D=0$ 时，$D_1, D_2, \cdots, D_n$ 如果不全都是零，(22)显然没有解；如果全部都是零，(22)也不一定有解. 但 $D=0$ 时，(23)是确有非零解的. 这将在下章详细地讨论.

克莱姆定理的意义是可以用方程组的系数及常数项借行列式把解明显地表达出来. 但实际引用时，需计算很多个行列式，尤其当 n 较大时，计算的工作量太大. 因此，在工程力学的实际问题中，解线性方程组一般不采用克莱姆规则，而采用逐次消去未知量的消去法，这方法就是中学代数中解 $n=2$ 或 $n=3$ 时的消元法. 这方法的实质就是把线性方程组化为系数的行列式是“三角形”的线性方程组，然后用逐次代入的方法解这三角形线性方程组. 所以说，克莱姆定理只是理论上的价值. 本书第二章论述线性方程组理论时，正是依据了克莱姆定理.

练 习 四

1．解方程组

$$\begin{cases} x+y+z=a+b+c, \\ ax+by+cz=a^2+b^2+c^2, \\ bcx+cay+abz=3abc, \end{cases}$$

式中 a, b, c 两两互异.

2．对于 $n+1$ 级行列式

$$0=\begin{vmatrix} a_{i1} & a_{i2} & \cdots & a_{in} & k_i \\ a_{11} & a_{12} & \cdots & a_{1n} & k_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & k_2 \\ \vdots & \vdots & & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} & k_n \end{vmatrix}.$$

试直接按它的第一行展开，来证明 $\left(\frac{D_1}{D}, \frac{D_2}{D}, \cdots, \frac{D_n}{D}\right)$ 为(22)之解.

3．利用克莱姆定理，证明：一个 n 次多项式不能有多于 n 个互异的零点.

第 5 节　拉普拉斯展开式、行列式的乘法

行列式可依它的任一行(列)展开的法则，可以推广到依它的任意若干行(列)来展开．行列式的这一展开法则，通常叫做拉普拉斯展开式，这正是本章开头提出要解决的第五个问题．

为此，先将余子式及代数余子式的概念给以推广．

在 n 级行列式

$$D=\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}$$

中，任意选取 k 个行与 k 个列(k 为任一自然数，且 $1\leqslant k\leqslant n$)，位置在这些行和列相交处的元素显然构成一个 k 级行列式，叫它为 D 的 *k 级子式*．也可以说，行列式 D 的 k 级子式是从 D 之行列中划掉 $n-k$ 个行和 $n-k$ 个列以后所剩下的一个 k 级行列式．特别，在 D 中划去一行和一列，得出一个 $n-1$ 级子式；而一级子式就是 D 的各个元素．显然，D 只有一个 n 级子式，就是 D 自身．

假定从 D 中取一个 k 级子式 M，划掉 M 所在的行与列后所剩下的一个 $n-k$ 级子式 M'，就叫做 M 的*余子式*．显然，划去 M' 所在的行和列就剩下了子式 M，于是，M 又是 M' 的余子式．因此，M 与 M' 是行列式 D 的(一对)*互余子式*．例如，元素 a_{ij} 和划去 D 之第 i 行及第 j 列所得的 $n-1$ 级子式组成一对互余子式．

如果 k 级子式 M 所在的行的序数为 i_1, i_2, $\cdots$, i_k，而所

在的列的序数为 $j_1, j_2, \cdots, j_k$，并令

$$s_M = i_1 + i_2 + \cdots + i_k + j_1 + j_2 + \cdots + j_k,$$

我们把 $(-1)^{s_M} \cdot M'$ 称为 M 的代数余子式，即

$$M \text{ 的代数余子式} = (-1)^{s_M} \times M \text{ 的余子式}.$$

例如，在 4 级行列式

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{vmatrix}$$

中，取第 2 行、第 3 行和第 1 列、第 3 列，就得到 2 级子式

$$M = \begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix},$$

它的余子式

$$M' = \begin{vmatrix} a_{12} & a_{14} \\ a_{42} & a_{44} \end{vmatrix},$$

代数余子式是

$$(-1)^{2+3+1+3} \begin{vmatrix} a_{12} & a_{14} \\ a_{42} & a_{44} \end{vmatrix} = -\begin{vmatrix} a_{12} & a_{14} \\ a_{42} & a_{44} \end{vmatrix}.$$

显然，这里介绍的余子式及代数余子式的概念是在第 3 节里所定义的余子式及代数余子式之概念的推广.

第 3 节中的引理也可以推广，即：

引理 *行列式 D 中任何一个 k 级子式与它的代数余子式的乘积，是行列式 D 中一部分的项.*

证明 先考虑 k 级子式 M 位于 D 之左上角的情况，即

$$D=\left|\begin{array}{ccc|ccc} a_{11} & \cdots & a_{1k} & a_{1,k+1} & \cdots & a_{1n} \\ \vdots & M & \vdots & \vdots & & \vdots \\ a_{k1} & \cdots & a_{kk} & a_{k,k+1} & \cdots & a_{kn} \\ \hline a_{k+1,1} & \cdots & a_{k+1,k} & a_{k+1,k+1} & \cdots & a_{k+1,n} \\ \vdots & & \vdots & \vdots & M' & \vdots \\ a_{n1} & \cdots & a_{nk} & a_{n,k+1} & \cdots & a_{nn} \end{array}\right|.$$

这时 M 的余子式 M' 就位于 D 的右下角. 因为 M 的行及列的序数同为 $1, 2, \cdots, k$, 所以 s_M 为偶数, 因此 M 的代数余子式就是它的余子式 M'. 引理的意义就是要证明乘积 MM' 的展开式是 D 中一部分的项.

因为 M 与 M' 分别为 k 级与 $n-k$ 级行列式, 由定义知

$$M=\sum_{(p_1,p_2,\cdots,p_k)}(-1)^{[p_1,p_2,\cdots,p_k]}a_{1p_1}a_{2p_2}\cdots a_{kp_k},$$

$$M'=\sum_{(p_{k+1},\cdots,p_n)}(-1)^{[p_{k+1},p_{k+2},\cdots,p_n]}a_{k+1,p_{k+1}}a_{k+2,p_{k+2}}\cdots a_{np_n},$$

式中 $(p_1, p_2, \cdots, p_k)$ 与 $(p_{k+1}, \cdots, p_n)$ 分别为 $(1, 2, \cdots, k)$ 与 $(k+1, k+2, \cdots, n)$ 的排列, $[p_1, p_2, \cdots, p_k]$ 与 $[p_{k+1}, p_{k+2}, \cdots, p_n]$ 分别为排列 $(p_1, p_2, \cdots, p_k)$ 与 $(p_{k+1}, p_{k+2}, \cdots, p_n)$ 的逆序数, 故

$$MM'=\sum(-1)^{[p_1,p_2,\cdots,p_k]+[p_{k+1},\cdots,p_n]}$$

$$\cdot a_{1p_1}a_{2p_2}\cdots a_{kp_k}a_{k+1,p_{k+1}}\cdots a_{np_n},$$

式中 $(p_1, p_2, \cdots, p_k, p_{k+1}, \cdots, p_n)$ 显为 $1, 2, \cdots, k, k+1, \cdots, n$ 的排列, 且因 $p_1, \cdots, p_k$ 是 $1, 2, \cdots, k$ 之排列而 $p_{k+1}, \cdots, p_n$ 是 $k+1, \cdots, n$ 之排列, 故任一 $p_i(i=1, \cdots, k)$ 与任一 $p_j(j=k+1, \cdots, n)$ 在排列 $(p_1, \cdots, p_k, p_{k+1}, \cdots, p_n)$ 中无逆序, 于是逆序数

$$[p_1, p_2, \cdots, p_k, p_{k+1}, \cdots, p_n]$$

$$=[p_1, p_k, \cdots, p_k]+[p_{k+1}, \cdots, p_n],$$

因而又得

$$MM' = \sum (-1)^{[p_1, \cdots, p_k, p_{k+1}, \cdots, p_n]} a_{1p_1} a_{2p_2} \cdots a_{kp_k} a_{k+1, p_{k+1}} \cdots a_{np_n},$$

式中 $p_1, p_2, \cdots, p_k$ 及 $p_{k+1}, \cdots, p_n$ 只分别为 1, 2, ⋯, k 及 $k+1, \cdots, n$ 的排列. 此即表明 MM' 中的每项都是 D 之项, 即 $MM' \subset D$, 且 MM' 包含了 D 之 $k!(n-k)!$ 个项.

再考虑一般情况, 即 D 之 k 级子式 M 位于 D 之第 i_1、i_2、⋯、i_k 行及第 j_1、j_2、⋯、j_k 列, 即

$$M = \begin{vmatrix} a_{i_1 j_1} & a_{i_1 j_2} & \cdots & a_{i_1 j_k} \\ a_{i_2 j_1} & a_{i_2 j_2} & \cdots & a_{i_2 j_k} \\ \vdots & \vdots & & \vdots \\ a_{i_k j_1} & a_{i_k j_2} & \cdots & a_{i_k j_k} \end{vmatrix}.$$

现在调动行列式 D 中行与列的位置, 使子式 M 搬到左上角, 且不使它的余子式变动. 这可以这样来进行: 即先把第 i_1 行与它前面的 i_1-1 个行一个挨一个地互换, 这样需要进行互换 i_1-1 回就能使第 i_1 行变为新行列式的第一行; 再同前面一样把第 i_2 行与它前面各行挨次互换, 直到它换到原来第 i_1 行的后面, 也就是换到第二行, 这样又互换了 i_2-2 回; 同样, 把第 i_3 行移到第三行去, 继续这样进行, 直到第 i_k 行移到第 k 行为止. 总起来, 共互换行的回数等于

$$\begin{aligned} &(i_1-1)+(i_2-2)+\cdots+(i_k-k) \\ &=(i_1+i_2+\cdots+i_k)-(1+2+\cdots+k). \end{aligned}$$

经过上述行的互换后, 子式 M 已经在新行列式的前 k 个行的上面. 接着, 再用上面的办法来挨次移动行列式的列: 先把第 j_1 列同它前面的列一一互换, 把它换到第一列, 再将第 j_2 列移到第二列, 继续这样进行, 直到第 j_k 列移到第 k 列, 总共对列移动了

$$(j_1+j_2+\cdots+j_k)-(1+2+\cdots+k)$$

回. 经过了上面行和列的变换后, 我们得出一个新的行列式

$\varDelta$，它的左上角即是 M. 因为上面这种行、列间的互换没有变动 D 的子式 M' 里面的行和列的相对位置，因此在 $\varDelta$ 中 M 的余子式仍然是 M'，但它已经在右下角了. 于是，根据上面证明的特殊情形，可知 $MM'\subset\varDelta$. 但由第 2 节的定理 2，又知道

$$D=(-1)^{(i_1+\cdots+i_k)+(j_1+\cdots+j_k)-2(1+2+\cdots+k)}\varDelta=(-1)^{s_M}\varDelta,$$

故
$$M\cdot(-1)^{s_M}M'\subset(-1)^{s_M}\varDelta=D.$$

这就证明了引理.

有了这个引理，就可以来证明下面的：

定理 1(拉普拉斯展开式)　*设在 n 级行列式*

$$D=\begin{vmatrix}a_{11} & \cdots & a_{1n}\\ \vdots & & \vdots\\ a_{n1} & \cdots & a_{nn}\end{vmatrix}$$

中任意取 k 个行(列)，那末含在这 k 个行(列)中所有 k 级子式与它们对应的代数余子式的乘积之和等于行列式 D.

证明　假设在 D 中所选定的 k 个行的序数为 i_1，i_2，…，$i_k(i_1<i_2<\cdots<i_k)$，由这 k 个行里能构成的所有 k 级子式用 M_1，M_2，…，M_t 表示，它们对应的代数余子式分别用 A_1，A_2，…，A_t 来表示，这里的 t，是由 n 个数中取 k 个的组合总数，即

$$t=C_n^k=\frac{n(n-1)\cdots(n-k+1)}{k!}.$$

下面就是要证明

$$D=M_1A_1+M_2A_2+\cdots+M_tA_t.$$

根据引理，可知每个积 $M_iA_i\subset D$.

再证任两个积 M_iA_i 与 $M_jA_j(i\neq j)$ 没有共同的项：这是因为，M_i 与 M_j 都是从 D 之第 i_1，i_2，…，i_k 行 $(i_1<i_2<\cdots<i_k)$ 里所取得的两个不同的 k 级行列式，故可设

$$M_i=\begin{vmatrix} a_{i_1j_1} & a_{i_1j_2} & \cdots & a_{i_1j_k} \\ a_{i_2j_1} & a_{i_2j_2} & \cdots & a_{i_2j_k} \\ \vdots & \vdots & & \vdots \\ a_{i_kj_1} & a_{i_kj_2} & \cdots & a_{i_kj_k} \end{vmatrix},$$

$$M_j=\begin{vmatrix} a_{i_1t_1} & a_{i_1t_2} & \cdots & a_{i_1t_k} \\ a_{i_2t_1} & a_{i_2t_2} & \cdots & a_{i_2t_k} \\ \vdots & \vdots & & \vdots \\ a_{i_kt_1} & a_{i_kt_2} & \cdots & a_{i_kt_k} \end{vmatrix},$$

虽然 $j_1, j_2, \cdots, j_k$ 及 $t_1, t_2, \cdots, t_k$ 都是从 1, 2, …, n 中所取的 k 个数，但 $j_1, j_2, \cdots, j_k$ 与 $t_1, t_2, \cdots, t_k$ 是不一数的，因而对于包含在 M_iA_i 内的 D 中诸项以及包含在 M_jA_j 内的 D 中诸项，因为作为在 D 之第 i_1、i_2、…、i_k 行之元所在的列分别是 $j_1, j_2, \cdots, j_k$ 及 $t_1, t_2, \cdots, t_k$ 之排列，故 M_iA_i 与 M_jA_j 没有共同的项.

因 $M_iA_i(i=1, \cdots, t)$ 含有行列式 D 之 $k!(n-k)!$ 个项，故 $M_1A_1+M_2A_2+\cdots+M_tA_t$ 共含有 D 之

$$t\cdot k!(n-k)!=C_n^k\cdot k!(n-k)!=n!$$

个不同的项，即包含了 D 的全部的项，于是有

$$D=M_1A_1+M_2A_2+\cdots+M_tA_t,$$

定理 1 完全得证.

据拉普拉斯展开式，容易得知两个 n 级行列式相乘的规则：设有两个 n 级行列式

$$D_1=\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} \text{ 与 } D_2=\begin{vmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & & \vdots \\ b_{n1} & b_{n2} & \cdots & b_{nn} \end{vmatrix},$$

据拉普拉斯展开式，有

$$\Delta = D_1 \cdot D_2 = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} & 0 & 0 & \cdots & 0 \\ a_{21} & a_{22} & \cdots & a_{2n} & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} & 0 & 0 & \cdots & 0 \\ x_{11} & x_{12} & \cdots & x_{1n} & b_{11} & b_{12} & \cdots & b_{1n} \\ x_{21} & x_{22} & \cdots & x_{2n} & b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{nn} & b_{n1} & b_{n2} & \cdots & b_{nn} \end{vmatrix},$$

式中 x_{ij} 可为任意的数. 为了便于看出 $D_1 \cdot D_2$ 能表为一个 n 级行列式,就令

$$x_{ij} = \delta_{ij} = \begin{cases} 1, & \text{当 } i = j \text{ 时}; \\ 0, & \text{当 } i \neq j \text{ 时}, \end{cases}$$

于是有

$$\Delta = D_1 D_2 = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} & 0 & 0 & \cdots & 0 \\ a_{21} & a_{22} & \cdots & a_{2n} & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} & 0 & 0 & \cdots & 0 \\ 1 & 0 & \cdots & 0 & b_{11} & b_{12} & \cdots & b_{1n} \\ 0 & 1 & \cdots & 0 & b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 1 & b_{n1} & b_{n2} & \cdots & b_{nn} \end{vmatrix}.$$

这时,用 $-b_{11}, -b_{21}, \cdots, -b_{n1}$ 分别乘 Δ 之第一、二、…、n 列,然后都加到 Δ 的第 $n+1$ 列上; 再用 $-b_{12}, -b_{22}, \cdots, -b_{n2}$ 分别乘 Δ 之第一、二、…、n 列后都加到 Δ 之第 $n+2$ 列上; 继续类似的手续,直到用 $-b_{1n}, -b_{2n}, \cdots, -b_{nn}$ 分别乘 Δ 之第一、二、…、n 列后都加到 Δ 的第 $2n$ 列上去. 这样,就可以得到

$$\Delta = D_1 D_2$$

$$= \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} & -\sum a_{1i}b_{i1} & -\sum a_{1i}b_{i2} & \cdots & -\sum a_{1i}b_{in} \\ a_{21} & a_{22} & \cdots & a_{2n} & -\sum a_{2i}b_{i1} & -\sum a_{2i}b_{i2} & \cdots & -\sum a_{2i}b_{in} \\ \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} & -\sum a_{ni}b_{i1} & -\sum a_{ni}b_{i2} & \cdots & -\sum a_{ni}b_{in} \\ 1 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 1 & 0 & 0 & \cdots & 0 \end{vmatrix},$$

再将上行列式按它的后 n 个列(即第 $n+1$、$n+2$、…、$2n$ 列)依拉普拉斯展开式展开，即得

$$D_1 D_2 = \begin{vmatrix} -\sum a_{1i}b_{i1} & -\sum a_{1i}b_{i2} & \cdots & -\sum a_{1i}b_{in} \\ -\sum a_{2i}b_{i1} & -\sum a_{2i}b_{i2} & \cdots & -\sum a_{2i}b_{in} \\ \vdots & \vdots & & \vdots \\ -\sum a_{ni}b_{i1} & -\sum a_{ni}b_{i2} & \cdots & -\sum a_{ni}b_{in} \end{vmatrix}$$

$$\times (-1)^{(n+1)+(n+2)+\cdots+2n+1+2+\cdots+n} \cdot 1$$

$$= (-1)^n \times \begin{vmatrix} \sum a_{1i}b_{i1} & \sum a_{1i}b_{i2} & \cdots & \sum a_{1i}b_{in} \\ \sum a_{2i}b_{i1} & \sum a_{2i}b_{i2} & \cdots & \sum a_{2i}b_{in} \\ \vdots & \vdots & & \vdots \\ \sum a_{ni}b_{i1} & \sum a_{ni}b_{i2} & \cdots & \sum a_{ni}b_{in} \end{vmatrix}$$

$$\times (-1)^{n(2n+1)}$$

$$= \begin{vmatrix} \sum a_{1i}b_{i1} & \sum a_{1i}b_{i2} & \cdots & \sum a_{1i}b_{in} \\ \sum a_{2i}b_{i1} & \sum a_{2i}b_{i2} & \cdots & \sum a_{2i}b_{in} \\ \vdots & \vdots & & \vdots \\ \sum a_{ni}b_{i1} & \sum a_{ni}b_{i2} & \cdots & \sum a_{ni}b_{in} \end{vmatrix}.$$

这就证得了

定理 2(行列式的乘法规则)　两个 n 级行列式

$$D_1=\begin{vmatrix}a_{11} & a_{12} & \cdots & a_{1n}\\ a_{21} & a_{22} & \cdots & a_{2n}\\ \vdots & \vdots & & \vdots\\ a_{n1} & a_{n2} & \cdots & a_{nn}\end{vmatrix} \text{ 与 } D_2=\begin{vmatrix}b_{11} & b_{12} & \cdots & b_{1n}\\ b_{21} & b_{22} & \cdots & b_{2n}\\ \vdots & \vdots & & \vdots\\ b_{n1} & b_{n2} & \cdots & b_{nn}\end{vmatrix}$$

的乘积 $D_1\cdot D_2$，可以表成这样一个 n 级行列式，它的第 i 行第 j 列的元素 c_{ij} 等于 D_1 之第 i 行的所有元素分别与 D_2 之第 j 列的对应元素的乘积之和，即

$$D_1D_2=C=\begin{vmatrix}c_{11} & c_{12} & \cdots & c_{1n}\\ c_{21} & c_{22} & \cdots & c_{2n}\\ \vdots & \vdots & & \vdots\\ c_{n1} & c_{n2} & \cdots & c_{nn}\end{vmatrix},$$

其中
$$c_{ij}=\sum_{t=1}^{n}a_{it}b_{tj}.$$

我们简称这个乘法规则为*行乘列*．由于任何行列式与它的转置行列式相等，即 $D_1=D_1'$，$D_2=D_2'$；且 D_i 之行、列又恰好是 D_i' 之列、行；于是，从 $C=D_1D_2=D_1D_2'=D_1'D_2=D_1'D_2'$ 可知，行列式的乘法规则还能归纳为：积 D_1D_2 之第 i 行第 j 列的元素 c_{ij} 为 D_1 之第 i 行之诸元与 D_2 之第 j 行诸对应元的积之和（从 D_1D_2' 来看），或可令为 D_1 之第 i 列之诸元与 D_2 之第 j 列诸对应元的积之和（从 $D_1'D_2$ 来看），或可令为 D_1 之第 i 列与 D_2 之第 j 行之诸对应元之积的和（从 $D_1'D_2'$ 来看）．所以，行列式的乘法规则共有四种，简称为：(i) *行乘列*，(ii) *行乘行*，(iii) *列乘列*，(iv) *列乘行*．使用时得根据具体情况．

［例 1］　计算对称行列式

$$D=\begin{vmatrix}1 & \cos(\alpha-\beta) & \cos(\alpha-\gamma)\\ \cos(\alpha-\beta) & 1 & \cos(\beta-\gamma)\\ \cos(\alpha-\gamma) & \cos(\beta-\gamma) & 1\end{vmatrix}.$$

解 $D=\begin{vmatrix}\cos(\alpha-\alpha) & \cos(\alpha-\beta) & \cos(\alpha-\gamma)\\ \cos(\beta-\alpha) & \cos(\beta-\beta) & \cos(\beta-\gamma)\\ \cos(\gamma-\alpha) & \cos(\gamma-\beta) & \cos(\gamma-\gamma)\end{vmatrix}$

$$=\begin{vmatrix}\cos\alpha & \sin\alpha & 0\\ \cos\beta & \sin\beta & 0\\ \cos\gamma & \sin\gamma & 0\end{vmatrix}\cdot\begin{vmatrix}\cos\alpha & \cos\beta & \cos\gamma\\ \sin\alpha & \sin\beta & \sin\gamma\\ 0 & 0 & 0\end{vmatrix}=0.$$

[例 2] 设 $x_1, x_2, \cdots, x_n$ 是 n 个两两互异的实数，并令 $S_i=x_1^i+x_2^i+\cdots+x_n^i$，试证

$$D=\begin{vmatrix}S_0 & S_1 & \cdots & S_{n-1}\\ S_1 & S_2 & \cdots & S_n\\ \vdots & \vdots & & \vdots\\ S_{n-1} & S_n & \cdots & S_{2n-2}\end{vmatrix}>0.$$

证明 $D=\begin{vmatrix}1 & 1 & \cdots & 1\\ x_1 & x_2 & \cdots & x_n\\ \vdots & \vdots & & \vdots\\ x_1^{n-1} & x_2^{n-1} & \cdots & x_n^{n-1}\end{vmatrix}\cdot\begin{vmatrix}1 & x_1 & \cdots & x_1^{n-1}\\ 1 & x_2 & \cdots & x_2^{n-1}\\ \vdots & \vdots & & \vdots\\ 1 & x_n & \cdots & x_n^{n-1}\end{vmatrix}$

$$=\prod_{1\le j<i\le n}(x_i-x_j)^2>0.$$

练　习　五

1．设 $D_1=\begin{vmatrix}a_{11} & a_{12} & \cdots & a_{1n}\\ a_{21} & a_{22} & \cdots & a_{2n}\\ \vdots & \vdots & & \vdots\\ a_{n1} & a_{n2} & \cdots & a_{nn}\end{vmatrix}$，$D_2=\begin{vmatrix}b_{11} & b_{12} & \cdots & b_{1n}\\ b_{21} & b_{22} & \cdots & b_{2n}\\ \vdots & \vdots & & \vdots\\ b_{n1} & b_{n2} & \cdots & b_{nn}\end{vmatrix}$，试利用第 2 节中行列式的性质，直接证明

$$D=\begin{vmatrix}\sum a_{1i}b_{i1} & \sum a_{1i}b_{i2} & \cdots & \sum a_{1i}b_{in}\\ \sum a_{2i}b_{i1} & \sum a_{2i}b_{i2} & \cdots & \sum a_{2i}b_{in}\\ \vdots & \vdots & & \vdots\\ \sum a_{ni}b_{i1} & \sum a_{ni}b_{i2} & \cdots & \sum a_{ni}b_{in}\end{vmatrix}=D_1\cdot D_2.$$

2. 求行列式

$$\Delta=\begin{vmatrix} a & b & c & d \\ b & a & d & c \\ c & d & a & b \\ d & c & b & a \end{vmatrix} \quad 与 \quad S=\begin{vmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & -1 & -1 & 1 \end{vmatrix}$$

的乘积,并从而计算 Δ。

3. 设 ω 是 1 的一个虚立方根,用拉普拉斯展开式,计算

$$\begin{vmatrix} 1 & 1 & 0 & 0 & 1 & 0 \\ 1 & \omega & 0 & 0 & \omega^2 & 0 \\ a_1 & b_1 & 1 & 1 & c_1 & 1 \\ a_2 & b_2 & 1 & \omega^2 & c_2 & \omega \\ a_3 & b_3 & 1 & \omega & c_3 & \omega^2 \\ 1 & \omega^2 & 0 & 0 & \omega & 0 \end{vmatrix}.$$

4. 设 $\varphi_1+\varphi_2+\varphi_3=0$,试证

$$\begin{vmatrix} 1 & \cos\varphi_3 & \cos\varphi_2 \\ \cos\varphi_3 & 1 & \cos\varphi_1 \\ \cos\varphi_2 & \cos\varphi_1 & 1 \end{vmatrix}=0.$$

5. 试证

$$\begin{vmatrix} a_{11} & \cdots & a_{1n} & b_{11} & \cdots & b_{1n} \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} & b_{n1} & \cdots & b_{nn} \\ \lambda & & 0 & \mu & & 0 \\ & \ddots & & & \ddots & \\ 0 & & \lambda & 0 & & \mu \end{vmatrix}=\begin{vmatrix} a_{11}\mu-b_{11}\lambda & \cdots & a_{1n}\mu-b_{1n}\lambda \\ \vdots & & \vdots \\ a_{n1}\mu-b_{n1}\lambda & \cdots & a_{nn}\mu-b_{nn}\lambda \end{vmatrix}.$$

第二章 线性方程组

在第一章里从解线性方程组的需要引进了行列式，最后又利用行列式去解线性方程组(克莱姆定理)，所以说克莱姆定理是第一章的重要问题．但是，引用克莱姆定理是有条件的，它要求线性方程组中方程的个数与未知量的个数相等，且由它们的系数所组成的行列式还要不等于零．在一些实际问题中，往往所碰到的线性方程组并不是这样简单的，有时候虽然方程组中方程的个数与未知量的个数相等，但由系数所组成的行列式却等于零；又有时候甚至是方程的个数不等于未知量的个数．对于这样一般的线性方程组，就不能象前章所说的那样，直接引用克莱姆定理来求解，问题就复杂了．

我们讨论线性方程组主要是讨论三个问题：

1．线性方程组究竟有没有解？即有解的条件是什么？

2．若有解，究竟有多少个解？又用什么方法去求解？

3．有解时，若解又不止一个，那末这许多解之间关系怎样？

我们先看几个例子，然后再归纳．

［例 1］ 解 $\begin{cases} 2x-y+3z=2, \\ x+2y-z=1. \end{cases}$

解 这时，方程的个数少于未知量的个数．因为

$$\begin{vmatrix} 2 & -1 \\ 1 & 2 \end{vmatrix} \neq 0,$$

所以把 z 看成已知数时，移到方程的右边，可用克莱姆定理得

$x=1-z,\ y=z$(而 z 得取任何值).

[例 2] 解 $\begin{cases} x+2y-z=2, \\ 2x-3y+z=3, \\ 4x+y-z=7. \end{cases}$

解 因这时

$$D=\begin{vmatrix} 1 & 2 & -1 \\ 2 & -3 & 1 \\ 4 & 1 & -1 \end{vmatrix}=0,$$

故不能直接引用克莱姆定理来求解. 如用 f_1, f_2, f_3 分别表示方程左边的多项式,考虑它们系数的关系,则发现有

$$f_3=2\times f_1+f_2,$$

同时,易知方程右边的常数项也有

$$7=2\times 2+3$$

的类似关系,于是第三个方程可由前两个方程(用 2 乘第一个方程然后相加)而推得. 因此,前两个方程的解也必为第三个方程的解,这时只需解前两个方程即可,即

$$\begin{cases} x+2y-z=2, \\ 2x-3y+z=3. \end{cases}$$

对于上方程组,可以采取象解例 1 一样的方法,用克莱姆定理,可得

$$x=\frac{1}{7}(z+12),\ \ y=\frac{1}{7}(3z+1),$$

而 z 取任意值.

[例 3] 解 $\begin{cases} f_1\equiv x-y+z=1, \\ f_2\equiv x+y-z=2, \\ f_3\equiv 5x-y+z=3. \end{cases}$

解 这时 $D=0$, 并易知

$$f_3=3f_1+2f_2, \quad 3\neq 3\times 1+2\times 2,$$

故方程组无解.

关于 n 个未知量 m 个方程组成的线性方程组，恒可假定 $m\leqslant n$(这是因为，如果 $m>n$，便可适当添加系数为零的新的未知量，使未知量的个数等于方程的个数．这样，由系数组成的 m 级行列式显然等于零，而把问题归结到如例 2 或例 3).

因而，从上面的例题可知：由未知量的系数组成的 m 级行列式假如不全都是零(象例 1 那样)，我们可用克莱姆定理去求方程组的解，当 $m<n$ 时它有无穷多个解；假如全部都是零，我们又能够找出方程左边它们间的关系，如果方程右边常数项也有同样的关系(象例 2 那样)，我们可删掉一些方程而只需解少于 m 个的方程，然后再进行类似的讨论；如果常数项象例 3 那样没有那种同样的关系，那末方程组就没有解．于是，讨论线性方程组时，如何决定方程左边与右边常数项有无那种同样关系(象例 2、3 中所说的)是一个主要问题．为此，下面第一节专门来讨论这个问题，从而引进一些必要的概念.

第 1 节　n 元向量、矩阵及其秩

我们知道，线性方程组是完全由它的系数决定，未知量的记号是不起什么作用的．因此，为了方便，从每一个线性方程把未知量分离出来，剩下的系数就成为一组有顺序的数．于是，讨论线性方程组就变成了讨论若干个有顺序的数组的问题，这就需要引进向量这一概念.

定义 1　把 n 个复数 $a_1, a_2, \cdots, a_n$ 排成一个有顺序的一组数

$$\boldsymbol{a}=(a_1, a_2, \cdots, a_n),$$

叫做 n 元向量，并把 $a_1, a_2, \cdots, a_n$ 叫做向量 $\boldsymbol{\alpha}$ 的分量。

有时，n 元向量又写成下面的形式

$$\begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}.$$

横写的叫行向量，直写的叫列向量．为节省书写篇幅，n 元向量一般都写为行向量的形式．

我们知道，在(二维)平面上，当坐标系取定后，由原点引出的向量可用两个有顺序的实数来确定．同样，在三维空间中，当坐标系取定后，由原点引出的向量可用三个有顺序的实数来确定．于是，上述 n 元向量的定义，也可以说是二维平面向量和三维空间向量的推广．不过，当坐标系取定后，两个有顺序的实数以及三个有顺序的实数可以分别看做平面上及三维空间中的向量，而当 $n>3$ 时，n 个实数所成的 n 元向量已没有直觉的几何意义，所谓"n 元向量"，不过是借用几何的术语而已．但是，n 元向量这一概念，还是有它的现实意义的．譬如确定三维空间中一个球面，就只需要知道球心的位置及半径的大小，假如把球心的坐标及半径的长按一定的顺序写出，就得到一个 4 元向量．所以，n 元向量这一概念，同数学上其他概念一样，是舍弃了现实对象的一个抽象．于是，我们把一个 n 个未知量的线性方程的系数组成一个 n 元向量，这线性方程的任一个解也可看做一个 n 元向量，叫做它的解向量．

两个线性方程，如果它们对应项的系数相等，它们本身就一致，因此可以叫两个 n 元向量 $\boldsymbol{\alpha}=(a_1, a_2, \cdots, a_n)$ 与 $\boldsymbol{\beta}=(b_1, b_2, \cdots, b_n)$ 是相等的，记为 $\boldsymbol{\alpha}=\boldsymbol{\beta}$，当 $a_i=b_i$ 时 $(i=1,$

$2, \cdots, n$). 这与平面上或三维空间中两个向量的对应分量相等从而这两向量就相等是一致的.

因为两个线性方程的和的系数，是它们对应项系数的和；一个数与一个线性方程的乘积的系数，是这个数与方程系数的乘积，于是加以抽象，可以定义 n 元向量的运算如下：

定义 2 设向量 $\boldsymbol{\alpha}=(a_1, a_2, \cdots, a_n)$，$\boldsymbol{\beta}=(b_1, b_2, \cdots, b_n)$，则把 $(a_1+b_1, a_2+b_2, \cdots, a_n+b_n)$ 称为 $\boldsymbol{\alpha}$ 与 $\boldsymbol{\beta}$ 的**和**，记成 $\boldsymbol{\alpha}+\boldsymbol{\beta}$，即

$$\boldsymbol{\alpha}+\boldsymbol{\beta}=(a_1+b_1, a_2+b_2, \cdots, a_n+b_n);$$

又，当 k 为数时，把 $k\boldsymbol{\alpha}=\boldsymbol{\alpha}k=(ka_1, ka_2, \cdots, ka_n)$ 称为**数 k 与向量 $\boldsymbol{\alpha}$ 的乘积**.

我们知道，平面上或三维空间中两个向量的和，是根据平行四边形的规则来演算的；如果用它的分量来表示，就是：两个向量的和的分量，等于这两个向量对应分量的和. 至于数 k 与向量 $\boldsymbol{\alpha}$ 的积 $k\boldsymbol{\alpha}$ 的意义，是指：当 $k>0$ 时，把 $\boldsymbol{\alpha}$ 拉长 k 倍；当 $k<0$ 时把 $\boldsymbol{\alpha}$ 拉长 $|k|$ 倍，并且方向变成与它相反；如果用分量来表示，就是：$k\boldsymbol{\alpha}$ 的分量等于 k 与 $\boldsymbol{\alpha}$ 的分量的积. 因此，上面的规定与平面上或三维空间中向量的加法及数与向量的乘法是一致的.

由定义易知

$$\boldsymbol{\alpha}+\boldsymbol{\beta}=\boldsymbol{\beta}+\boldsymbol{\alpha},\ (\boldsymbol{\alpha}+\boldsymbol{\beta})+\boldsymbol{\gamma}=\boldsymbol{\alpha}+(\boldsymbol{\beta}+\boldsymbol{\gamma}),$$

这是说，**向量的加法满足交换律及结合律**. 因为

$$(a_1, \cdots, a_n)+(0, \cdots, 0)=(a_1, \cdots, a_n),$$

这是说，向量 $(0, \cdots, 0)$ 与向量 $\boldsymbol{\alpha}=(a_1, \cdots, a_n)$ 的和仍然是 $\boldsymbol{\alpha}$，所以 $(0, \cdots, 0)$ 有象数 0 相仿的作用，因此叫它为**零向量**，即 $\mathbf{0}=(0, 0, \cdots, 0)$.

要注意，零向量记号 $\mathbf{0}$ 与数 0 容易混淆，因此，在一般书

写时，如果不便于区分粗体字，同一记号在不同的地方就有不同的解释，例如，如果向量间等式的一边是零，那末这个零应该就是零向量.

因为$(a_1, a_2, \cdots, a_n)+(-a_1, -a_2, \cdots, -a_n)=(0, 0, \cdots, 0)$，所以，把$(-a_1, -a_2, \cdots, -a_n)$称为$\boldsymbol{\alpha}=(a_1, a_2, \cdots, a_n)$的负向量，用$-\boldsymbol{\alpha}$表示，即规定$(-1)\boldsymbol{\alpha}=-\boldsymbol{\alpha}$.

我们把$\boldsymbol{\alpha}+(-\boldsymbol{\beta})$写成$\boldsymbol{\alpha}-\boldsymbol{\beta}$，叫做$\boldsymbol{\alpha}$与$\boldsymbol{\beta}$的差，即

$$\boldsymbol{\alpha}-\boldsymbol{\beta}=(a_1-b_1, a_2-b_2, \cdots, a_n-b_n).$$

显然，$(\boldsymbol{\alpha}-\boldsymbol{\beta})+\boldsymbol{\beta}=\boldsymbol{\alpha}$，因此向量的加法有逆运算，也就是说，向量还有减法运算.

由定义可知:

$$k(\boldsymbol{\alpha}\pm\boldsymbol{\beta})=k\boldsymbol{\alpha}\pm k\boldsymbol{\beta}, \quad (k\pm l)\boldsymbol{\alpha}=k\boldsymbol{\alpha}\pm l\boldsymbol{\alpha},$$
$$k(l\boldsymbol{\alpha})=(kl)\boldsymbol{\alpha}, \qquad 1\cdot\boldsymbol{\alpha}=\boldsymbol{\alpha}.$$

此外，又有

$$0\boldsymbol{\alpha}=\boldsymbol{0}, \quad k\boldsymbol{0}=\boldsymbol{0}, \quad (-k)\boldsymbol{\alpha}=k(-\boldsymbol{\alpha})=-k\boldsymbol{\alpha}.$$

再如，由$k\boldsymbol{\alpha}=\boldsymbol{0}$，我们可推出$k=0$或$\boldsymbol{\alpha}=\boldsymbol{0}$(这是因为，如果$k\neq0$，用$\frac{1}{k}$乘$k\boldsymbol{\alpha}=\boldsymbol{0}$的两边，可得到

$$\frac{1}{k}(k\boldsymbol{\alpha})=\boldsymbol{0},$$
$$\left(\frac{1}{k}\cdot k\right)\boldsymbol{\alpha}=\boldsymbol{0},$$
$$1\cdot\boldsymbol{\alpha}=\boldsymbol{0},$$
$$\boldsymbol{\alpha}=\boldsymbol{0}.)$$

前面所提出的方程间的那种关系，若用向量来表达，我们就有

定义 3 设$\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_m$是$m$个向量，如果存在不全为零的数$k_1, k_2, \cdots, k_m$，使得

$$k_1\boldsymbol{\alpha}_1+k_2\boldsymbol{\alpha}_2+\cdots+k_m\boldsymbol{\alpha}_m=\mathbf{0}, \tag{1}$$

则 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_m$ 叫做线性相关．如果象这样的数 $k_1, k_2, \cdots, k_m$ 不存在，换句话说，欲使(1)成立，必有 $k_1=k_2=\cdots=k_m=0$（即 $k_1, k_2, \cdots, k_m$ 必都为零），则就叫 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_m$ 线性无关．

譬如，例 2 中的向量

$$\boldsymbol{\alpha}_1=(1, 2, -1, 2), \quad \boldsymbol{\alpha}_2=(2, -3, 1, 3),$$
$$\boldsymbol{\alpha}_3=(4, 1, -1, 7)$$

是线性相关的，因为 $2\boldsymbol{\alpha}_1+\boldsymbol{\alpha}_2-\boldsymbol{\alpha}_3=\mathbf{0}$．但 $\boldsymbol{\alpha}=(1, 2)$ 与 $\boldsymbol{\beta}=(3, 4)$线性无关，因为从 $k\boldsymbol{\alpha}+l\boldsymbol{\beta}=\mathbf{0}$ 即得 $k+3l=0$, $2k+4l=0$, 所以 $k=l=0$．

当 $m=2$ 时，(1)式就成为 $k_1\boldsymbol{\alpha}_1+k_2\boldsymbol{\alpha}_2=\mathbf{0}$．因为 k_1、k_2 不完全是零，故若 $k_1\neq0$，就有 $\boldsymbol{\alpha}_1=-\frac{k_2}{k_1}\boldsymbol{\alpha}_2$，这时又可说 $\boldsymbol{\alpha}_1$, $\boldsymbol{\alpha}_2$ 成比例．因此，对两个向量来说，“线性相关”与“成比例”这两个概念是一致的．

假如 $k_1, \cdots, k_m$ 适合(1)式，显然，与它们成比例的数 $kk_1, \cdots, kk_m$ 也适合(1)式，这里 k 是任意数．但要注意的是：当 $\boldsymbol{\alpha}_1, \cdots, \boldsymbol{\alpha}_m$ 线性相关时，适合(1)式的数 $k_1, \cdots, k_m$，除与这一组数成比例的数外，还可能有与它们不成比例的，也就是说，它们的线性关系一般不是唯一的．譬如四个三元向量

$$\boldsymbol{\alpha}_1=(5, 2, 1), \quad \boldsymbol{\alpha}_2=(-1, 3, 3),$$
$$\boldsymbol{\alpha}_3=(9, 7, 5), \quad \boldsymbol{\alpha}_4=(3, 8, 7)$$

的线性相关的关系有

$$4\boldsymbol{\alpha}_1-\boldsymbol{\alpha}_2-3\boldsymbol{\alpha}_3+2\boldsymbol{\alpha}_4=\mathbf{0};$$

也还有

$$2\boldsymbol{\alpha}_1+\boldsymbol{\alpha}_2-\boldsymbol{\alpha}_3+0\boldsymbol{\alpha}_4=\mathbf{0},$$
$$0\boldsymbol{\alpha}_1+3\boldsymbol{\alpha}_2+\boldsymbol{\alpha}_3-2\boldsymbol{\alpha}_4=\mathbf{0}.$$

注意：两个向量只有当它们的元数相同，才能有相等或不相等的意义，也只有在这时，才能有和、差的运算．只有当一组向量的元数相等时才能讨论它们的线性相关或是线性无关．对于元数不同的向量，来讨论相等与否，或是求它们的和或差，或是向量组的线性相关与否，都是没有意义的．总之，今后凡是牵涉到向量间的某种关系问题时，总是假定它们的元数都相同．

向量间的线性相关与无关，还与另一个概念是紧密相连的，即向量间的线性组合．

定义 4 假定 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_m$ 是 m 个向量，那末 $k_1\boldsymbol{\alpha}_1+k_2\boldsymbol{\alpha}_2+\cdots+k_m\boldsymbol{\alpha}_m$ 叫做 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_m$ 的线性组合．如果向量

$$\boldsymbol{\alpha}=k_1\boldsymbol{\alpha}_1+k_2\boldsymbol{\alpha}_2+\cdots+k_m\boldsymbol{\alpha}_m,$$

则 $\boldsymbol{\alpha}$ 也叫做 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_m$ 的线性组合，或者叫 $\boldsymbol{\alpha}$ 可以由 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_m$ 来线性表示．

关于线性相关和线性组合，有下面的：

性质 1 m 个向量 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_m$ 线性相关的充要条件，是其中有某一个向量为其余 $m-1$ 个向量的线性组合．

证明 先证必要性．从 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_m$ 线性相关，知有不完全为 0 的 m 个数 $k_1, \cdots, k_m$，使 $k_1\boldsymbol{\alpha}_1+\cdots+k_m\boldsymbol{\alpha}_m=\mathbf{0}$．既然 $k_1, \cdots, k_m$ 不全为零，必至少有一不等于零，不妨设 $k_m\neq 0$，即可得到

$$\boldsymbol{\alpha}_m=-\frac{k_1}{k_m}\boldsymbol{\alpha}_1-\cdots-\frac{k_{m-1}}{k_m}\boldsymbol{\alpha}_{m-1},$$

即 $\boldsymbol{\alpha}_m$ 为其余 $m-1$ 个 $\boldsymbol{\alpha}_1, \cdots, \boldsymbol{\alpha}_{m-1}$ 的线性组合．

再证充分性．若 $\boldsymbol{\alpha}_m$ 为 $\boldsymbol{\alpha}_1, \cdots, \boldsymbol{\alpha}_{m-1}$ 的线性组合：

$$\boldsymbol{\alpha}_m=c_1\boldsymbol{\alpha}_1+\cdots+c_{m-1}\boldsymbol{\alpha}_{m-1},$$

则有 $$c_1\boldsymbol{\alpha}_1+\cdots+c_{m-1}\boldsymbol{\alpha}_{m-1}+(-1)\boldsymbol{\alpha}_m=\mathbf{0},$$

这表明 $\boldsymbol{\alpha}_1, \cdots, \boldsymbol{\alpha}_m$ 线性相关.

性质 2 任意一个非零向量是线性无关的. 又，m 个向量 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_m$ 中，当有一个为零向量时，这 m 个向量恒为线性相关. $\boldsymbol{\alpha}_1, \cdots, \boldsymbol{\alpha}_m$ 中若有 k 个 $(k<m)$ 线性相关，则 $\boldsymbol{\alpha}_1, \cdots, \boldsymbol{\alpha}_m$ 也必线性相关.

证明 1) 从 $k\boldsymbol{\alpha}=\mathbf{0}$ 且 $\boldsymbol{\alpha}\neq\mathbf{0}$ 时，显然有 $k=0$，这说明了任一个非零向量线性无关.

2) 若 $\boldsymbol{\alpha}_1=\mathbf{0}$，则 $1\cdot\boldsymbol{\alpha}_1+0\cdot\boldsymbol{\alpha}_2+\cdots+0\cdot\boldsymbol{\alpha}_m=\mathbf{0}$，此即表明 $\boldsymbol{\alpha}_1, \cdots, \boldsymbol{\alpha}_m$ 线性相关.

3) 若 $c_1\boldsymbol{\alpha}_1+\cdots+c_k\boldsymbol{\alpha}_k=\mathbf{0}$($k<m$，且 $c_1, \cdots, c_k$ 不全为零)，则 $c_1\boldsymbol{\alpha}_1+\cdots+c_k\boldsymbol{\alpha}_k+0\boldsymbol{\alpha}_{k+1}+\cdots+0\boldsymbol{\alpha}_m=\mathbf{0}$，显然，这表明 $\boldsymbol{\alpha}_1, \cdots, \boldsymbol{\alpha}_m$ 线性相关.

在引言里所指出的线性方程组的诸方程间的那种关系，我们已可用向量间的线性相关、无关及线性组合的概念来表述了. 例 2 还说明：当方程组的诸方程间有线性相关的那种关系时，我们可删掉多余的方程而只需考虑保留的方程. 所谓从方程组中删掉一些方程保留一些方程，这无异乎是说从一个向量组中怎样挑选出一组线性无关的向量，使其余向量都可表为这一组向量的线性组合. 要解决这问题，就要判定方程组之各方程之系数所成之向量组是线性相关还是线性无关. 为此，又需引用一个重要的工具——矩阵.

定义 5 mn 个数 $a_{ij}(i=1, \cdots, m;\ j=1, \cdots, n)$ 排成 m 行、n 列的表

$$\boldsymbol{A}=\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \cdots & \cdots & \cdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}$$

叫做 $m\times n$-矩阵，或简称矩阵，也可简记为 $\boldsymbol{A}=(a_{ij})$. a_{ij} 叫

做 A 中第 i 行、第 j 列上的数或元．当 $m=n$ 时，A 又叫做 n 级矩阵，也叫做 n 级方阵．

注意：矩阵与行列式是两个完全不同的概念，n 级行列式是一个数，而 n 级矩阵不是一个数，而是 n^2 个有顺序的数(或元)的表．不要由于它们的表示形式差不多而产生混淆．

由 $m\times n$–矩阵 A 中任意一行的 n 个数而成的 n 元向量，叫做 A 的行向量；由 A 中任意一列的 m 个数而成的 m 元向量，叫做 A 的列向量．因此 A 有 m 个行向量，n 个列向量．一般，行向量横写，列向量竖写，A 的第 i 个行向量写成

$$(a_{i1},\ a_{i2},\ \cdots,\ a_{in}),$$

第 j 个列向量写成

$$\begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{nj} \end{pmatrix}.$$

只有一行或一列的矩阵，又常叫做向量．

同行列式的情形一样，在 $m\times n$–矩阵 A 中，取某 k 个行及某 k 个列($k\leqslant m,\ n$)，由这些行、列相交处的元构成的 k 级行列式叫做 A 的 k 级子式．显然，n 级矩阵只有一个 n 级子式，这子式又常叫做矩阵的行列式．n 级矩阵 A 的行列式用 $\det A$ 表示，或简记为 $|A|$．

例如，在 2×3–矩阵

$$A=\begin{pmatrix} 1 & 2 & -1 \\ 2 & -3 & 1 \end{pmatrix}$$

中，1 级子式是由其中一个元构成的，因此共有六个 1 级子式．它的三个 2 级子式是

$$\begin{vmatrix}1 & 2\\2 & -3\end{vmatrix},\quad\begin{vmatrix}1 & -1\\2 & 1\end{vmatrix},\quad\begin{vmatrix}2 & -1\\-3 & 1\end{vmatrix}.$$

关于矩阵，有一个重要概念，即：

定义 6 矩阵 $\boldsymbol{A}$ 中不等于零的子式的最高级数如果是 r，我们就叫 $\boldsymbol{A}$ 的秩是 r. 如果 n 级矩阵的秩是 n（即它的行列式不等于零），就叫做满秩矩阵，否则叫降秩矩阵.

性质 3 秩是零的矩阵是全由数零组成的矩阵. 又，矩阵 $\boldsymbol{A}$ 中若至少有一个 r 级子式不为零，并且所有 $r+1$ 级子式都为零，那末 $\boldsymbol{A}$ 的秩为 r.

证明 所有 $r+1$ 级子式都为零时，根据第一章第 5 节拉普拉斯展开式可知，高于 $r+1$ 级的子式显然也都是零，故 $\boldsymbol{A}$ 中不等于零的子式的最高级数为 r，这就证得了 $\boldsymbol{A}$ 的秩是 r.

关于矩阵的行（列）向量组，怎样判别是线性相关或无关的呢？对此，有下面的：

定理 1 $m\times n$-矩阵 $\boldsymbol{A}$ 的 m 个行向量线性相关的充要条件是 $\boldsymbol{A}$ 的秩小于 m.

证明 先证必要性. 设 $m\times n$-矩阵 $\boldsymbol{A}=(a_{ij})$ 的 m 个行向量

$$\boldsymbol{\alpha}_1=(a_{11},\ a_{12},\ \cdots,\ a_{1n}),\ \boldsymbol{\alpha}_2=(a_{21},\ a_{22},\ \cdots,\ a_{2n}),\ \cdots,$$
$$\boldsymbol{\alpha}_m=(a_{m1},\ a_{m2},\ \cdots,\ a_{mn})$$

线性相关，则由性质 1，这 m 个行向量中有某一行向量是其余 $m-1$ 个向量的线性组合，不妨假定 $\boldsymbol{\alpha}_m$ 是 $\boldsymbol{\alpha}_1,\ \cdots,\ \boldsymbol{\alpha}_{m-1}$ 的线性组合，即 $\boldsymbol{\alpha}_m=k_1\boldsymbol{\alpha}_1+\cdots+k_{m-1}\boldsymbol{\alpha}_{m-1}$. 因此，用 $-k_1,\ \cdots,\ -k_{m-1}$ 分别乘 $\boldsymbol{A}$ 的第 1 行，…，第 $m-1$ 行，然后都加到第 m 行上，就使第 m 行中所有元都为零. 所以，如果 $\boldsymbol{A}$ 有 m 级子式，那末 $\boldsymbol{A}$ 的任意一个 m 级子式都是零，这就表明 $\boldsymbol{A}$ 的秩小于 m，定理的必要条件得证.

再证充分性．设 $\boldsymbol{A}$ 的秩为 r，$r<m$．若 $r=0$，则每一 $\boldsymbol{\alpha}_i$ 为零向量，故 $\boldsymbol{\alpha}_1, \cdots, \boldsymbol{\alpha}_m$ 线性相关．若 $r>0$，不失一般性，可假定 $\boldsymbol{A}$ 之左上角的 r 级子式 $D\neq 0$，即

$$D=\begin{vmatrix} a_{11} & \cdots & a_{1r} \\ \vdots & & \vdots \\ a_{r1} & \cdots & a_{rr} \end{vmatrix}\neq 0.$$

如果我们能够证明 $r+1$ 个向量 $\boldsymbol{\alpha}_1, \cdots, \boldsymbol{\alpha}_{r+1}$ 线性相关，则据性质 2，即得 $\boldsymbol{\alpha}_1, \cdots, \boldsymbol{\alpha}_m$ 也线性相关，条件的充分性就解决了．于是下面即是要证 $\boldsymbol{\alpha}_1, \cdots, \boldsymbol{\alpha}_{r+1}$ 线性相关．为此，要找不完全为零的 $r+1$ 个数 $k_1, \cdots, k_{r+1}$，使

$$\begin{aligned} & k_1\boldsymbol{\alpha}_1+k_2\boldsymbol{\alpha}_2+\cdots+k_{r+1}\boldsymbol{\alpha}_{r+1} \\ & \quad =(k_1a_{11}+k_2a_{21}+\cdots+k_{r+1}a_{r+1,1}, \\ & \qquad \cdots, k_1a_{1n}+k_2a_{2n}+\cdots+k_{r+1}a_{r+1,n}) \\ & \quad =\boldsymbol{0}, \end{aligned}$$

也就是

$$k_1a_{1t}+k_2a_{2t}+\cdots+k_{r+1}a_{r+1,t}=0. \quad (t=1, 2, \cdots, n)$$

为了求出这 $r+1$ 个数 $k_1, k_2, \cdots, k_{r+1}$，我们考虑下面 $r+1$ 级行列式

$$D_t=\begin{vmatrix} a_{11} & \cdots & a_{1r} & a_{1t} \\ \vdots & & \vdots & \vdots \\ a_{r1} & \cdots & a_{rr} & a_{rt} \\ a_{r+1,1} & \cdots & a_{r+1,r} & a_{r+1,t} \end{vmatrix}.$$

当 $t>r$（假如这样的 t 存在）时，D_t 为 $\boldsymbol{A}$ 的一个 $r+1$ 级子式，所以 $D_t=0$．然将 D_t 按最后一列展开，得

$$A_1a_{1t}+\cdots+A_ra_{rt}+Da_{r+1,t}=0, \tag{2}$$

这里 $A_1, \cdots, A_r, D(\neq 0)$ 分别是 $a_{1t}, \cdots, a_{rt}, a_{r+1,t}$ 在 D_t 中的代数余子式．显然，$A_1, \cdots, A_r$ 与 t 无关．当 $t\leqslant r$ 时，

D_t 有两列相同，所以也有 $D_t=0$，这时(2)式同样成立．这就是说，对于 $t=1, 2, \cdots, n$，总有 $D_t=0$，即(2)式都成立．于是，$A_1, \cdots, A_r, D$ 就是我们要求的 $r+1$ 个不完全为零的数，所以 $\boldsymbol{\alpha}_1, \cdots, \boldsymbol{\alpha}_{r+1}$ 线性相关，即定理的充分条件成立．

至此，定理 1 完全得证．

同理，有：

定理 1′ $m\times n$-矩阵 $\boldsymbol{A}$ 的 n 个列向量线性相关的充要条件是 $\boldsymbol{A}$ 的秩小于 n．

由定理 1，即可得到：

推论 任意 $m(>n)$ 个 n 元向量线性相关．又，n 级行列式 $D=0$ 的充要条件是 n 个行(或列)都线性相关．

对于一组向量，其中线性无关的最多的个数有多大呢？对此，有：

定理 2 设一组向量的集合中含有 r 个向量，使得该集合中任一向量都为这 r 个向量的线性组合，则这集合中任意 $r+1$ 个向量都必线性相关．

证明 在集合中任取 $r+1$ 个向量：

$$\boldsymbol{\beta}_1=(b_{11}, b_{12}, \cdots, b_{1n}),$$
$$\boldsymbol{\beta}_2=(b_{21}, b_{22}, \cdots, b_{2n}),$$
$$\cdots\cdots\cdots\cdots\cdots\cdots$$
$$\boldsymbol{\beta}_r=(b_{r1}, b_{r2}, \cdots, b_{rn}),$$
$$\boldsymbol{\beta}_{r+1}=(b_{r+1,1}, b_{r+1,2}, \cdots, b_{r+1,n}).$$

由题设，每个 $\boldsymbol{\beta}_i$ 都可表为

$$\boldsymbol{\alpha}_1=(a_{11}, a_{12}, \cdots, a_{1n}),$$
$$\boldsymbol{\alpha}_2=(a_{21}, a_{22}, \cdots, a_{2n}),$$
$$\cdots\cdots\cdots\cdots\cdots\cdots,$$
$$\boldsymbol{\alpha}_r=(a_{r1}, a_{r2}, \cdots, a_{rn})$$

的线性组合，即

$$\boldsymbol{\beta}_1=\lambda_1'\boldsymbol{\alpha}_1+\lambda_2'\boldsymbol{\alpha}_2+\cdots+\lambda_r'\boldsymbol{\alpha}_r,$$

$$\boldsymbol{\beta}_2=\lambda_1''\boldsymbol{\alpha}_1+\lambda_2''\boldsymbol{\alpha}_2+\cdots+\lambda_r''\boldsymbol{\alpha}_r,$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

$$\boldsymbol{\beta}_{r+1}=\lambda_1^{(r+1)}\boldsymbol{\alpha}_1+\lambda_2^{(r+1)}\boldsymbol{\alpha}_2+\cdots+\lambda_r^{(r+1)}\boldsymbol{\alpha}_r.$$

现在考虑矩阵

$$\boldsymbol{B}=\begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & & \vdots \\ b_{r1} & b_{r2} & \cdots & b_{rn} \\ b_{r+1,1} & b_{r+1,2} & \cdots & b_{r+1,n} \end{pmatrix}$$

中的 $r+1$ 级子式(如果存在，即 $r+1\leqslant n$)．不失普遍性，令其中任一个 $r+1$ 级子式为

$$D_{r+1}=\begin{vmatrix} b_{11} & b_{12} & \cdots & b_{1r} & b_{1,r+1} \\ b_{21} & b_{22} & \cdots & b_{2r} & b_{2,r+1} \\ \vdots & \vdots & & \vdots & \vdots \\ b_{r1} & b_{r2} & \cdots & b_{rr} & b_{r,r+1} \\ b_{r+1,1} & b_{r+1,2} & \cdots & b_{r+1,r} & b_{r+1,r+1} \end{vmatrix}.$$

于是，根据第一章第 2 节，可把 D_{r+1} 写成若干行列式之和，如

$$D_{r+1}=\begin{vmatrix} \sum\limits_{t=1}^{r}\lambda_t'a_{t1} & \sum\limits_{t=1}^{r}\lambda_t'a_{t2} & \cdots & \sum\limits_{t=1}^{r}\lambda_t'a_{tr} & \sum\limits_{t=1}^{r}\lambda_t'a_{t,r+1} \\ \vdots & \vdots & & \vdots & \vdots \\ \sum\limits_{t=1}^{r}\lambda_t^{(r)}a_{t1} & \sum\limits_{t=1}^{r}\lambda_t^{(r)}a_{t2} & \cdots & \sum\limits_{t=1}^{r}\lambda_t^{(r)}a_{tr} & \sum\limits_{t=1}^{r}\lambda_t^{(r)}a_{t,r+1} \\ \sum\limits_{t=1}^{r}\lambda_t^{(r+1)}a_{t1} & \sum\limits_{t=1}^{r}\lambda_t^{(r+1)}a_{t2} & \cdots & \sum\limits_{t=1}^{r}\lambda_t^{(r+1)}a_{tr} & \sum\limits_{t=1}^{r}\lambda_t^{(r+1)}a_{t,r+1} \end{vmatrix}$$

$$=\sum_{t_1=1}^{r}\sum_{t_2=1}^{r}\cdots\sum_{t_{r+1}=1}^{r}a_{t_1 1}a_{t_2 2}\cdots a_{t_r r}a_{t_{r+1}, r+1}$$

$$\cdot\begin{vmatrix}\lambda'_{t_1} & \lambda'_{t_2} & \cdots & \lambda'_{t_r} & \lambda'_{t_{r+1}}\\ \vdots & \vdots & & \vdots & \vdots\\ \lambda^{(r)}_{t_1} & \lambda^{(r)}_{t_2} & \cdots & \lambda^{(r)}_{t_r} & \lambda^{(r)}_{t_{r+1}}\\ \lambda^{(r+1)}_{t_1} & \lambda^{(r+1)}_{t_2} & \cdots & \lambda^{(r+1)}_{t_r} & \lambda^{(r+1)}_{t_{r+1}}\end{vmatrix};$$

但因每个 $t_i(i=1, 2, \cdots, r, r+1)$ 只能在 $1, 2, \cdots, r$ 中取值，故 $r+1$ 个值 $t_1, t_2, \cdots, t_r, t_{r+1}$ 中必有相同的，故不得不有

$$\begin{vmatrix}\lambda'_{t_1} & \lambda'_{t_2} & \cdots & \lambda'_{t_r} & \lambda'_{t_{r+1}}\\ \vdots & \vdots & & \vdots & \vdots\\ \lambda^{(r)}_{t_1} & \lambda^{(r)}_{t_2} & \cdots & \lambda^{(r)}_{t_r} & \lambda^{(r)}_{t_{r+1}}\\ \lambda^{(r+1)}_{t_1} & \lambda^{(r+1)}_{t_2} & \cdots & \lambda^{(r+1)}_{t_r} & \lambda^{(r+1)}_{t_{r+1}}\end{vmatrix}=0,$$

因此 $D_{r+1}=0$，此即表明矩阵 $\boldsymbol{B}$ 的秩 $<r+1$，据定理 1，可知 $\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \cdots, \boldsymbol{\beta}_r, \boldsymbol{\beta}_{r+1}$ 线性相关．这就证得了定理 2．

联系矩阵的秩与其行(列)向量间的关系，又有：

定理 3 矩阵 $\boldsymbol{A}$ 的秩等于 r 的充要条件是 $\boldsymbol{A}$ 中有 r 个行(列)向量线性无关，并且任意 $r+1$ 个行(列)向量都线性相关．

证明 先证必要性．设 $\boldsymbol{A}$ 的秩为 r．于是 $\boldsymbol{A}$ 中有一个 r 级子式不等于零，而任意的 $r+1$ 级子式都为零．据定理 1 可知，不为零的 r 级子式所在的 r 个行(列)向量必线性无关，并且任意 $r+1$ 个行(列)向量都线性相关．这就证得了必要性．

再证充分性．设矩阵 $\boldsymbol{A}$ 中已有 r 个行(列)向量线性无关，且任意 $r+1$ 个行(列)向量都线性相关．那末，根据定理 1(1′)得知，在所与的 r 个线性无关的行(列)向量中，必有一个 r 级子式不为零，且由任 $r+1$ 个行(列)向量中取出的任

$r+1$ 级子式都必为零，这说明了 $\boldsymbol{A}$ 中任意的 $r+1$ 级子式都为零．由性质 3，知 $\boldsymbol{A}$ 之秩为 r.

至此，定理 3 证毕．

由定理 3，即得：

推论 1 *矩阵的秩、行向量组里面线性无关的最大数目以及列向量组中线性无关的最大数目，这三者是一致的．*

再利用定理 2，又得判别矩阵之秩的：

推论 2 *设矩阵 $\boldsymbol{A}$ 中有一个 r 级子式不等于零，但凡包含它为子式的 $\boldsymbol{A}$ 中所有 $r+1$ 级子式都为零，那末 $\boldsymbol{A}$ 的秩是 r.* （注意这里的推论 2 与前述性质 3 的差异）

证明 不损一般性，可令矩阵

$$\boldsymbol{A}=\begin{pmatrix} a_{11} & \cdots & a_{1r} & \cdots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ a_{r1} & \cdots & a_{rr} & \cdots & a_{rn} \\ \vdots & & \vdots & & \vdots \\ a_{m1} & \cdots & a_{mr} & \cdots & a_{mn} \end{pmatrix}$$

中一个不为零的 r 级子式在 $\boldsymbol{A}$ 之左上角，即

$$D_r=\begin{vmatrix} a_{11} & \cdots & a_{1r} \\ \vdots & & \vdots \\ a_{r1} & \cdots & a_{rr} \end{vmatrix}\neq 0.$$

设矩阵 $\boldsymbol{A}$ 之 m 个行向量为

$$\begin{aligned} \boldsymbol{\alpha}_1 &= (a_{11}, \cdots, a_{1n}), \\ \boldsymbol{\alpha}_2 &= (a_{21}, \cdots, a_{2n}), \\ &\cdots\cdots\cdots\cdots\cdots\cdots, \\ \boldsymbol{\alpha}_m &= (a_{m1}, \cdots, a_{mn}). \end{aligned}$$

讨论以 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_r$ 及任一 $\boldsymbol{\alpha}_i(i=r+1, \cdots, m)$ 共 $r+1$ 个向量为行向量的 $\boldsymbol{A}$ 之子矩阵（$(r+1)\times n$-矩阵）

$$B=\begin{pmatrix}a_{11} & \cdots & a_{1r} & \cdots & a_{1n}\\ \vdots & & \vdots & & \vdots\\ a_{r1} & \cdots & a_{rr} & \cdots & a_{rn}\\ a_{i1} & \cdots & a_{ir} & \cdots & a_{in}\end{pmatrix},$$

由题设 $\begin{vmatrix}a_{11} & \cdots & a_{1r} & a_{1t}\\ \vdots & & \vdots & \vdots\\ a_{r1} & \cdots & a_{rr} & a_{rt}\\ a_{i1} & \cdots & a_{ir} & a_{it}\end{vmatrix}=0$，$(t=r+1,\ \cdots,\ n)$ 就可知道

$$C=\begin{vmatrix}a_{11} & \cdots & a_{1r} & a_{1t}\\ \vdots & & \vdots & \vdots\\ a_{r1} & \cdots & a_{rr} & a_{rt}\\ a_{i1} & \cdots & a_{ir} & a_{it}\end{vmatrix}=0,\ (t=1,\cdots,r,r+1,\cdots,n)$$

这是因为，当 $t=1,\cdots,r$ 时，C 中有两个列一致，所以可把下标 t 从 $t=r+1,\ \cdots,\ n$ 扩展到 $t=1,\ \cdots,\ n$. 但是，按最后一列展开 C，又得

$$A_1a_{1t}+\cdots+A_ra_{rt}+D_ra_{it}=0,\ (t=1,\ \cdots,\ r,\ r+1,\ \cdots,\ n)$$

其中 $A_1,\ \cdots,\ A_r,\ D_r$ 是 $a_{1t},\ \cdots,\ a_{rt},\ a_{it}$ 在 C 内的代数余子式，显然都与 t 无关. 由此即得

$$A_1\boldsymbol{\alpha}_1+\cdots+A_r\boldsymbol{\alpha}_r+D_r\boldsymbol{\alpha}_i=\mathbf{0},$$

故由 $D_r\neq0$ 得知 $\boldsymbol{\alpha}_i$ 为 $\boldsymbol{\alpha}_1,\ \cdots,\ \boldsymbol{\alpha}_r$ 的线性组合. 即 $\boldsymbol{\alpha}_1,\ \cdots,\ \boldsymbol{\alpha}_r,\ \boldsymbol{\alpha}_{r+1},\ \cdots,\ \boldsymbol{\alpha}_m$ 中任一个都是前 r 个的线性组合，据定理 2，可知 $\boldsymbol{A}$ 中任 $r+1$ 个行向量线性相关. 但根据定理 1，由 $D_r\neq0$ 又保证了 $\boldsymbol{A}$ 有 r 个行向量 $\boldsymbol{\alpha}_1,\ \cdots,\ \boldsymbol{\alpha}_r$ 线性无关. 于是，根据定理 3，知 $\boldsymbol{A}$ 的秩为 r，证完.

练 习 六

1. 假定 $\boldsymbol{\alpha}_1,\ \cdots,\ \boldsymbol{\alpha}_m$ 是元数相同的 m 个向量. 如果存在 m 个全为零的数 $k_1,\ \cdots,\ k_m$，使 $k_1\boldsymbol{\alpha}_1+\cdots+k_m\boldsymbol{\alpha}_m=\mathbf{0}$，问 $\boldsymbol{\alpha}_1,\ \cdots,\ \boldsymbol{\alpha}_m$ 是否线性无关？如果任何不

全为零的 m 个数 $k_1, \cdots, k_m$，使 $k_1\boldsymbol{\alpha}_1+\cdots+k_m\boldsymbol{\alpha}_m\neq\mathbf{0}$，问 $\boldsymbol{\alpha}_1, \cdots, \boldsymbol{\alpha}_m$ 是否线性无关？

2．设 $\boldsymbol{\alpha}_1, \cdots, \boldsymbol{\alpha}_m, \boldsymbol{\beta}_1, \cdots, \boldsymbol{\beta}_m$ 是元数相同的一组向量，如果仅当 $k_1=\cdots=k_m=0$ 时，才成立 $k_1\boldsymbol{\alpha}_1+\cdots+k_m\boldsymbol{\alpha}_m+k_1\boldsymbol{\beta}_1+\cdots+k_m\boldsymbol{\beta}_m=\mathbf{0}$，问 $\boldsymbol{\alpha}_1, \cdots, \boldsymbol{\alpha}_m, \boldsymbol{\beta}_1, \cdots, \boldsymbol{\beta}_m$ 是否线性无关？

3．设有元数相同的向量组

(I) $\boldsymbol{\alpha}_1, \cdots, \boldsymbol{\alpha}_r$;

(II) $\boldsymbol{\alpha}_1, \cdots, \boldsymbol{\alpha}_r, \boldsymbol{\alpha}_{r+1}, \cdots, \boldsymbol{\alpha}_m$.

问：当(I)无关时，(II)无关吗？又，当(II)相关时，(I)相关吗？

4．设有两个向量组

$$\boldsymbol{\alpha}_1=(a_{11}, \cdots, a_{1r}), \cdots, \boldsymbol{\alpha}_m=(a_{m1}, \cdots, a_{mr});$$

$$\boldsymbol{\beta}_1=(a_{11},\cdots,a_{1r},a_{1,r+1},\cdots,a_{1n}), \cdots, \boldsymbol{\beta}_m=(a_{m1},\cdots,a_{mr},a_{m,r+1},\cdots,a_{mn}).$$

试证：如果 $\boldsymbol{\alpha}_1, \cdots, \boldsymbol{\alpha}_m$ 线性无关，那末 $\boldsymbol{\beta}_1, \cdots, \boldsymbol{\beta}_m$ 也线性无关．它的逆是否成立？为什么？

5．在秩是 r 的矩阵中，有没有等于零的 $r-1$ 级子式？有没有等于零的 r 级子式？有没有不等于零的 $r+1$ 级子式？

6．试证：$m\times n$-矩阵 (a_{ij}) 的秩为 0 或 1 的充要条件是存在 $m+n$ 个数 $\lambda_1, \cdots, \lambda_m, \mu_1, \cdots, \mu_n$，使得 $a_{ij}=\lambda_i\mu_j$，$(i=1, \cdots, m; j=1, \cdots, n)$.

第 2 节　线性方程组的解的讨论

有了前节的准备，就可以着手解决线性方程组的问题．本节只讨论在引言里提出的第 1、2 两个问题，即解决：线性方程组是否有解，以及当它有解时，解究竟有多少？

对于线性方程组

$$\begin{cases} a_{11}x_1+a_{12}x_2+\cdots+a_{1n}x_n=b_1, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{m1}x_1+a_{m2}x_2+\cdots+a_{mn}x_n=b_m. \end{cases} \tag{3}$$

作两个矩阵，一为 $m\times n$-矩阵，另一为 $m\times(n+1)$-矩阵，分别为

$$\boldsymbol{A}=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \cdots\cdots & \cdots\cdots & \cdots\cdots & \cdots\cdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$$

与
$$\boldsymbol{B}=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ \cdots\cdots & \cdots\cdots & \cdots\cdots & \cdots\cdots & \cdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{pmatrix},$$

分别叫做方程组(3)的系数矩阵及增广矩阵. $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 的秩分别记为 r_A 与 r_B.

先假定方程组(3)有解，即有一组数(x_1, x_2, $\cdots$, x_n)满足(3). 于是，对增广矩阵 B 说来，意味着 $\boldsymbol{B}$ 的最后一列为前 n 列(即 A 之 n 个列)的线性组合，故在 $\boldsymbol{B}$ 之列向量组中找它们线性无关的最大数目就等于在 $\boldsymbol{A}$ 之列向量组中线性无关的最大数目，这说明了 $r_B=r_A$.

反过来，假定 $r_B=r_A$，并设它们都为 r，即 $r_B=r_A=r$. 不损一般性，可设

$$\begin{vmatrix} a_{11} & \cdots & a_{1r} \\ \vdots & & \vdots \\ a_{r1} & \cdots & a_{rr} \end{vmatrix} \neq 0.$$

于是，因
$$\boldsymbol{C}=\begin{pmatrix} a_{11} & \cdots & a_{1r} & b_1 \\ \vdots & & \vdots & \vdots \\ a_{r1} & \cdots & a_{rr} & b_r \\ \vdots & & \vdots & \vdots \\ a_{m1} & \cdots & a_{mr} & b_m \end{pmatrix},$$

故 $r_C \leqslant r_B=r$; 另一方面，由前节的定理 1′，得知 $\boldsymbol{C}$ 之前 r 个列线性无关，而由前节定理 8 的推论 1，得 $r_C \geqslant r$; 这就证得了 $r_C=r$. 因而$\begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}$为 $\boldsymbol{C}$ 之前 r 个列的线性组合，当然也是 $\boldsymbol{A}$

之所有 n 个列的线性组合，此即表明(3)有解.

通过上面的讨论，就证得了：

定理 1 *线性方程组(3)有解的充要条件是系数矩阵与增广矩阵有相等的秩.*

定理 1 已回答了引言里提出的第 1 个问题，即线性方程组有解的条件.

再来解决引言里提出的第 2 个问题，即当(3)有解时，(3)之解数有多少？又怎样去求解？

设(3)有解，即 $r_A=r_B$ 并记为 r. 不失一般性，假定

$$\begin{vmatrix} a_{11} & \cdots & a_{1r} \\ \vdots & & \vdots \\ a_{r1} & \cdots & a_{rr} \end{vmatrix} \neq 0.$$

于是，取(3)之前 r 个方程以及后面 $m-r$ 个方程中的任一个方程(假定存在的话，即 $m>r$)，共 $r+1$ 个方程，如

$$\begin{cases} a_{11}x_1+\cdots+a_{1r}x_r+a_{1,r+1}x_{r+1}+\cdots+a_{1n}x_n=b_1, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{r1}x_1+\cdots+a_{rr}x_r+a_{r,r+1}x_{r+1}+\cdots+a_{rn}x_n=b_r, \\ a_{i1}x_1+\cdots+a_{ir}x_r+a_{i,r+1}x_{r+1}+\cdots+a_{in}x_n=b_i, \end{cases} \tag{4}$$

这里 $i=r+1, \cdots, m$. 根据前节的定理 1，矩阵

$$\begin{pmatrix} a_{11} & \cdots & a_{1r} & a_{1,r+1} & \cdots & a_{1n} & b_1 \\ \vdots & & \vdots & \vdots & & \vdots & \vdots \\ a_{r1} & \cdots & a_{rr} & a_{r,r+1} & \cdots & a_{rn} & b_r \\ a_{i1} & \cdots & a_{ir} & a_{i,r+1} & \cdots & a_{in} & b_i \end{pmatrix}$$

的前 r 个行线性无关而其全部的 $r+1$ 个行又线性相关，故最后一行为前 r 个行的线性组合. 这说明了(4)中前 r 个方程的解必为最后一方程的解，也就是说，(3)之前 r 个方程的解必为其余方程之解. 所以，欲解(3)，就只需解其前 r 个方程即可. 而(3)之前 r 个方程又可改写为

$$a_{11}x_1+\cdots+a_{1r}x_r=b_1-a_{1,r+1}x_{r+1}-\cdots-a_{1n}x_n,$$
$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$
$$a_{r1}x_1+\cdots+a_{rr}x_r=b_r-a_{r,r+1}x_{r+1}-\cdots-a_{rn}x_n.$$

于是，据克莱姆定理，可知 $x_1,\cdots,x_r$ 可唯一地表为 $x_{r+1},\cdots,x_n$ 的线性表示式，而 $x_{r+1},\cdots,x_n$ 均得取任意数. 这不仅说明了当(3)有解时怎样去求解，而且还知道：在 $r=n$ 时(3)有唯一的一组解，在 $r<n$ 时(3)有无限多组解(因这时至少有一个未知量得取任意值). 也就是证得了：

定理 2 线性方程组(3)如果有解(即 $r_A=r_B=r$)，那末，当矩阵之秩 $r=n$ 时有唯一组解，当 $r<n$ 时有无限多组解. (3)的解都能用下述方法可求得：取 $\boldsymbol{A}$ 中任一个不等于零的 r 级子式所在的 r 个方程，并将相应于这个 r 级子式的未知量当做未知量，而将其余的未知量(当它们存在，即 $r<n$ 时)视为常量，移到所取的 r 个方程的右边，然后再利用克莱姆定理解这 r 个未知量 r 个方程所组成的方程组，就能得到(3)的全部解.

定理 2 已回答了引言内提出的第 2 个问题，也就是当线性方程组有解时，怎样去求解以及解的个数.

由定理 2，还可知道：当线性方程组有解时，解要么就只有一组，要么就有无穷多组. 不会出现解不止一组而又只有有穷多组的情况.

练 习 七

1. 问线性方程组

$$\begin{cases}\lambda x_1+x_2+x_3=1,\\ x_1+\lambda x_2+x_3=\lambda,\\ x_1+x_2+\lambda x_3=\lambda^2,\end{cases}$$

当 λ 取何值时无解，λ 取何值时有唯一解，λ 取何值时有无穷多解？

2．试证：方程组

$$\begin{aligned} x_1-x_2&=a_1, \\ x_2-x_3&=a_2, \\ x_3-x_4&=a_3, \\ &\cdots\cdots\cdots\cdots, \\ x_{n-1}-x_n&=a_{n-1}, \\ x_n-x_1&=a_n \end{aligned}$$

有解的充要条件是 $\sum_{i=1}^{n} a_i=0$；并且，在有解情形下，求出它的一般解。

3．设线性方程组

$$\begin{cases} a_{11}x_1+\cdots+a_{1n}x_n=b_1, \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ a_{n1}x_1+\cdots+a_{nn}x_n=b_n \end{cases}$$

之系数矩阵的秩等于矩阵

$$\begin{pmatrix} a_{11} & \cdots & a_{1n} & b_1 \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & \cdots & a_{nn} & b_n \\ b_1 & \cdots & b_n & 0 \end{pmatrix}$$

的秩，试证方程组有解。

4．讨论空间三个平面

$$\begin{aligned} a_1x+b_1y+c_1z&=d_1, \\ a_2x+b_2y+c_2z&=d_2, \\ a_3x+b_3y+c_3z&=d_3 \end{aligned}$$

的位置关系。

5．试证明：平面中三条不同的直线

$$\begin{aligned} ax+by+c&=0, \\ bx+cy+a&=0, \\ cx+ay+b&=0 \end{aligned}$$

相交于一点的充要条件是 $a+b+c=0$。

第3节 基础解系

前已指出，对于线性方程组(3)，当 $r_A=r_B=r<n$ 时，有无限多组解. 本节就是要研究，这无穷多组解之间究竟有怎样的关系？也就是要回答引言里提出的第3个问题.

我们易知：对于所有 n 元向量之集合 $\mathfrak{B}$，其中线性无关的向量的最大数目是多少呢？我们知道，有如下 n 个向量：

$$\begin{aligned}\boldsymbol{e}_1&=(1,\ 0,\ 0,\ \cdots,\ 0,\ 0),\\ \boldsymbol{e}_2&=(0,\ 1,\ 0,\ \cdots,\ 0,\ 0),\\ &\cdots\cdots\cdots\cdots\cdots\cdots,\\ \boldsymbol{e}_n&=(0,\ 0,\ 0,\ \cdots,\ 0,\ 1),\end{aligned}$$

它们显然线性无关，且任一个 n 元向量 $\boldsymbol{a}=(c_1,\ c_2,\ \cdots,\ c_n)$ 又为它们的线性组合，即

$$\boldsymbol{a}=\sum_{i=1}^{n}c_i\boldsymbol{e}_i.$$

于是，据第57页的定理2，$\mathfrak{B}$ 中任 $n+1$ 个向量必线性相关.

对于方程组(3)之解向量之集合 $\mathfrak{S}$，当然是 $\mathfrak{B}$ 的子集合，故 $\mathfrak{S}$ 中任 $n+1$ 个向量也必线性相关. 这就是说，(3)之解向量之集合中，线性无关的向量之最大数目不会超过 n. 今问到底是多少呢？

为解决这问题，先考虑简单的情况，即齐次方程组：

$$\begin{cases}a_{11}x_1+\cdots+a_{1n}x_n=0,\\ \cdots\cdots\cdots\cdots\cdots\cdots\\ a_{m1}x_1+\cdots+a_{mn}x_n=0.\end{cases}\tag{5}$$

设(5)之系数矩阵

$$\boldsymbol{A}=\begin{pmatrix}a_{11} & \cdots & a_{1n}\\ \cdots & \cdots & \cdots\\ a_{m1} & \cdots & a_{mn}\end{pmatrix}$$

的秩 $r_A=r<n$，则令

$$\begin{vmatrix} a_{11} & \cdots & a_{1r} \\ \cdots & \cdots & \cdots \\ a_{r1} & \cdots & a_{rr} \end{vmatrix} \neq 0$$

时，据前节的理论，(5)之全部解可由(5)之前 r 个方程

$$\begin{cases} a_{11}x_1+\cdots+a_{1r}x_r=-a_{1,r+1}x_{r+1}-a_{1,r+2}x_{r+2}-\cdots-a_{1n}x_n, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{r1}x_1+\cdots+a_{rr}x_r=-a_{r,r+1}x_{r+1}-a_{r,r+2}x_{r+2}-\cdots-a_{rn}x_n \end{cases} \tag{6}$$

根据克莱姆定理而求得(其中 $x_{r+1}, x_{r+2}, \cdots, x_n$ 得为任意值)．

为求(5)之解向量集合中线性无关向量之最大数目，不能令 $x_{r+1}, x_{r+2}, \cdots, x_n$ 都等于零，这是因为，当 $x_{r+1}=x_{r+2}=\cdots=x_n=0$ 时，由(6)即得 $x_1=\cdots=x_r=0$，随而得出(5)之零解 $(0, 0, \cdots, 0)$，这一解不是线性无关的．于是，为求(5)之线性无关的解，就必须令 $x_{r+1}, x_{r+2}, \cdots, x_n$ 不全为零．然对(5)之任一解向量 $(x_1, \cdots, x_r, \cdots, x_n)$，$x_1, \cdots, x_r$ 都可以表为 $x_{r+1}, x_{r+2}, \cdots, x_n$ 的线性表示式，如

$$\begin{cases} x_1=c_{1,r+1}x_{r+1}+c_{1,r+2}x_{r+2}+\cdots+c_{1n}x_n, \\ x_2=c_{2,r+1}x_{r+1}+c_{2,r+2}x_{r+2}+\cdots+c_{2n}x_n, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ x_r=c_{r,r+1}x_{r+1}+c_{r,r+2}x_{r+2}+\cdots+c_{rn}x_n. \end{cases} \tag{7}$$

为简单计，分别令 $(x_{r+1}, x_{r+2}, \cdots, x_n)$ 取 $n-r$ 组值为 $(1, 0, \cdots, 0)$，$(0, 1, 0, \cdots, 0)$，$\cdots$，$(0, 0, \cdots, 0, 1)$，则由(7)式就得到了(5)之 $n-r$ 个解向量：

$$\alpha_1=(c_{1,r+1}, c_{2,r+1}, \cdots, c_{r,r+1}, 1, 0, 0, \cdots, 0, 0),$$
$$\alpha_2=(c_{1,r+2}, c_{2,r+2}, \cdots, c_{r,r+2}, 0, 1, 0, \cdots, 0, 0),$$
$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$
$$\alpha_{n-r}=(c_{1n}, c_{2n}, \cdots, c_{rn}, 0, 0, 0, \cdots, 0, 1).$$

显然，由于把 $\alpha_1, \alpha_2, \cdots, \alpha_{n-r}$ 当做行向量而成的 $(n-r)\times n$-矩阵的最后 $n-r$ 个列之行列式

$$\begin{vmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 1 \end{vmatrix} \neq 0,$$

故据第 1 节的定理 1，可知 $\alpha_1, \alpha_2, \cdots, \alpha_{n-r}$ 线性无关. 而(5)之任意解向量 $(x_1, x_2, \cdots, x_r, x_{r+1}, x_{r+2}, \cdots, x_n)$ 又确为 $\alpha_1, \alpha_2, \cdots, \alpha_{n-r}$ 的线性组合：这是因为

$$\begin{aligned}&(x_1, \cdots, x_r, x_{r+1}, \cdots, x_n)\\ &=\left(\sum_{j=r+1}^{n} c_{1j}x_j, \cdots, \sum_{j=r+1}^{n} c_{rj}x_j, x_{r+1}, \cdots, x_n\right)\\ &=x_{r+1}\alpha_1+x_{r+2}\alpha_2+\cdots+x_n\alpha_{n-r}.\end{aligned}$$

于是，根据第 1 节的定理 2，得知(5)中任意 $n-r+1$ 个解向量是线性相关的. 这就证得了：

定理 1 当线性齐次方程组(5)之系数矩阵 A 之秩为 r 时，则(5)之解向量集合中线性无关解向量之最大个数等于 $n-r$. 也就是说，(5)之全部解向量中可找得 $n-r$ 个线性无关的向量，而任意 $n-r+1$ 个都线性相关.

定理 1 中的 $n-r$ 个线性无关的向量，叫做(5)之**基础解系**，一般地，有：

定义 1 线性方程组的解向量集合中，如有 k 个解向量具有下述两个性质：

i）这 k 个解向量线性无关；

ii）方程组之任意的解又为这 k 个解的线性组合，

就叫这 k 个解为方程组之**基础解系**.

于是，定理 1 又可改述为：齐次方程组(5)的基础解系

中，包含有$n-r$个线性无关的解(r为(5)之系数矩阵的秩)．

进一步，讨论(5)中k个解为(5)之基础解系的条件是什么？设$\boldsymbol{\beta}_1$，$\boldsymbol{\beta}_2$，…，$\boldsymbol{\beta}_k$为(5)之基础解系，当然，据定义知它们是线性无关的．若$k>n-r$，则与定理1中结论"任$n-r+1$个都是线性相关的"相矛盾，故k不能比$n-r$大；若$k<n-r$，由于(5)之任意解必为$\boldsymbol{\beta}_1$，…，$\boldsymbol{\beta}_k$之线性组合，故上面的$\boldsymbol{\alpha}_1$，…，$\boldsymbol{\alpha}_{n-r}$也应是$\boldsymbol{\beta}_1$，…，$\boldsymbol{\beta}_k$的线性组合，由第1节定理2，知$\boldsymbol{\alpha}_1$，…，$\boldsymbol{\alpha}_{n-r}$应线性相关，这又显非所许，故$k$也不能小于$n-r$．这就证明了$k=n-r$．反之，(5)中任$n-r$个线性无关的向量$\boldsymbol{\beta}_1$，…，$\boldsymbol{\beta}_{n-r}$必为(5)之基础解系．这是因为，取(5)之任一解$\boldsymbol{\gamma}$，则由定理1知$n-r+1$个解$\boldsymbol{\gamma}$，$\boldsymbol{\beta}_1$，…，$\boldsymbol{\beta}_{n-r}$线性相关，故由$\boldsymbol{\beta}_1$，…，$\boldsymbol{\beta}_{n-r}$之线性无关性得知$\boldsymbol{\gamma}$为$\boldsymbol{\beta}_1$，…，$\boldsymbol{\beta}_{n-r}$的线性组合，这就证明了$\boldsymbol{\beta}_1$，…，$\boldsymbol{\beta}_{n-r}$为(5)之基础解系．

再回顾上面形成$\boldsymbol{\alpha}_1$，…，$\boldsymbol{\alpha}_{n-r}$的方法，是使$x_{r+1}$，$x_{r+2}$，…，$x_n$分别为异于零的$n-r$级行列式

$$\begin{vmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 1 \end{vmatrix}$$

之第1、2、…、$n-r$行．然只要令x_{r+1}，x_{r+2}，…，x_n分别为任意一个异于零的$n-r$级行列式之第1、2、…、$n-r$行，而由(7)所决定(5)之$n-r$个解向量也是线性无关的，随而也必是(5)的基础解系．于是得知(5)有无穷多组基础解系．

这就证得了：

定理2 *线性齐次方程组(5)必有基础解系．当(5)之系数矩阵之秩为r时，每组基础解系必包含$n-r$个线性无关的解向量；且(5)含有无穷多组基础解系；而(5)中k个线性无关*

的解为基础解系的充要条件是 $k=n-r$. 又(5)之全部解，可由一基础解系之一切线性组合而得到.

现在再讨论非齐次的情况，即(3)中 $b_1, \cdots, b_m$ 不全为零. 这时，应在 $r_B=r_A(=r)$ 时(3)才有解. 且由前节的讨论得知：若

$$\begin{vmatrix} a_{11} & \cdots & a_{1r} \\ \vdots & & \vdots \\ a_{r1} & \cdots & a_{rr} \end{vmatrix} \neq 0$$

时，则只须解前 r 个方程即可. 也就是说，

$$\begin{cases} a_{11}x_1+\cdots+a_{1r}x_r=b_1-a_{1,r+1}x_{r+1}-\cdots-a_{1n}x_n, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{r1}x_1+\cdots+a_{rr}x_r=b_r-a_{r,r+1}x_{r+1}-\cdots-a_{rn}x_n \end{cases} \tag{8}$$

之解集合是与(3)之解集合一致. 然解(8)时，由克莱姆定理可知 $x_1, \cdots, x_r$ 都可以表为 $x_{r+1}, x_{r+2}, \cdots, x_n$ 的线性表示式，如

$$\begin{cases} x_1=\lambda_1+\lambda_{1,r+1}x_{r+1}+\lambda_{1,r+2}x_{r+2}+\cdots+\lambda_{1n}x_n, \\ x_2=\lambda_2+\lambda_{2,r+1}x_{r+1}+\lambda_{2,r+2}x_{r+2}+\cdots+\lambda_{2n}x_n, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ x_r=\lambda_r+\lambda_{r,r+1}x_{r+1}+\lambda_{r,r+2}x_{r+2}+\cdots+\lambda_{rn}x_n. \end{cases} \tag{9}$$

再令可自由取值的变量 $x_{r+1}, x_{r+2}, \cdots, x_n$ 分别为 $(1, 0, \cdots, 0)$, $(0, 1, \cdots, 0)$, $\cdots$, $(0, \cdots, 0, 1)$时，仍如上述理由，可由(9)得到 $n-r$ 个解

$$\begin{aligned} \boldsymbol{\alpha}_1&=(\lambda_1+\lambda_{1,r+1}, \lambda_2+\lambda_{2,r+1}, \cdots, \lambda_r+\lambda_{r,r+1}, 1, 0, 0, \cdots, 0, 0), \\ \boldsymbol{\alpha}_2&=(\lambda_1+\lambda_{1,r+2}, \lambda_2+\lambda_{2,r+2}, \cdots, \lambda_r+\lambda_{r,r+2}, 0, 1, 0, \cdots, 0, 0), \\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ \boldsymbol{\alpha}_{n-r}&=(\lambda_1+\lambda_{1n}, \lambda_2+\lambda_{2n}, \cdots, \lambda_r+\lambda_{rn}, 0, 0, 0, \cdots, 0, 1). \end{aligned}$$

但令 $x_{r+1}=\cdots=x_n=0$ 代入(9)中所得的解 $(\lambda_1, \lambda_2, \cdots, \lambda_r, 0, \cdots, 0)$ 也是一组非零解(即至少有一个 $\lambda_i\neq 0$). 这样，除了

上述的 $\alpha_i(i=1, 2, \cdots, n-r)$外，连同

$$\alpha_0=(\lambda_1, \lambda_2, \cdots, \lambda_r, 0, 0, 0, \cdots, 0, 0),$$

共得到了(8)的 $n-r+1$ 个解 $\alpha_i(i=1, 2, \cdots, n-r, 0)$，它们还是线性无关的(因为它们是一个 $(n-r+1)\times n$-矩阵的 $n-r+1$ 个行向量，且含有一 $n-r+1$ 级子式

$$\begin{vmatrix} \lambda_i+\lambda_{i,r+1} & 1 & 0 & 0 & \cdots & 0 & 0 \\ \lambda_i+\lambda_{i,r+2} & 0 & 1 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots \\ \lambda_i+\lambda_{in} & 0 & 0 & 0 & \cdots & 0 & 1 \\ \lambda_i & 0 & 0 & 0 & \cdots & 0 & 0 \end{vmatrix}\neq 0,$$

这里 $\lambda_i\neq 0$，据第1节的定理1，即知它们线性无关)。但(8)之任一个解 $(x_1, \cdots, x_r, x_{r+1}, \cdots, x_m)$，又为 α_i $(i=0, 1, 2, \cdots, n-r)$的线性组合：这是因为，据(9)式，有

$$\begin{aligned}&(x_1, x_2, \cdots, x_r, x_{r+1}, x_{r+2}, \cdots, x_n)\\ &=\left(\lambda_1+\sum_{j=r+1}^{n}\lambda_{1j}x_j,\quad \lambda_2+\sum_{j=r+1}^{n}\lambda_{2j}x_j, \cdots,\right.\\ &\qquad\left.\lambda_r+\sum_{j=r+1}^{n}\lambda_{rj}x_j,\quad x_{r+1},\quad x_{r+2}, \cdots, x_n\right)\\ &=x_{r+1}\alpha_1+x_{r+2}\alpha_2+\cdots+x_n\alpha_{n-r}\\ &\quad+(1-x_{r+1}-x_{r+2}-\cdots-x_n)\alpha_0.\end{aligned}$$

这说明了(8)有 $n-r+1$ 个线性无关的解，且任意解又为它们的线性组合。据定义1，可知(8)有基础解系。

注意：(8)之任意解 $(x_1, \cdots, x_n)$ 表为 $\alpha_i(i=0, 1, 2, \cdots, n-r)$之线性组合时，其组合表示式中之常数为 $x_{r+1}, x_{r+2}, \cdots, x_n$, $1-x_{r+1}-x_{r+2}-\cdots-x_n$，而其和恰等于1. 反之，当 $\sum_{i=0}^{n-r}c_i=1$ 时，易证 $\sum_{i=0}^{n-r}c_i\alpha_i$ 也是(8)的解(参看后面练习八的第3题)。于是，与证定理2一样，可知有：

定理3 含 n 个未知量的线性非齐次方程组，当其系数矩阵与增广矩阵有相等的秩 r 时，必有基础解系；每组基础解系包含 $n-r+1$ 个线性无关的向量；且它有无穷多组基础解系，而它的 k 个线性无关的解向量为基础解系的充要条件是 $k=n-r+1$．又，它的全部解可由一基础解系之一切线性组合使组合系数之和等于1时而得到．

定理2与定理3分别解决了线性齐次方程组与非齐次方程组的解之间关系．下面再来讨论非齐次方程组与相应的齐次方程组之间的解向量的相互关系．

设有非齐次线性方程组

$$\begin{cases} a_{11}x_1+\cdots+a_{1n}x_n=b_1, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{m1}x_1+\cdots+a_{mn}x_n=b_m, \end{cases} \tag{3}$$

系数矩阵 $\boldsymbol{A}$ 与增广矩阵 $\boldsymbol{B}$ 的秩相等($r_A=r_B$)；作相应的齐次方程组

$$\begin{cases} a_{11}x_1+\cdots+a_{1n}x_n=0, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{m1}x_1+\cdots+a_{mn}x_n=0. \end{cases} \tag{5}$$

令 $\boldsymbol{\gamma}_1, \boldsymbol{\gamma}_2, \cdots, \boldsymbol{\gamma}_{n-r}(r=r_A=r_B)$ 为(5)之基础解系，$\boldsymbol{\alpha}$ 为(3)之某一解．

若 $\boldsymbol{\beta}$ 为(3)之任一解，则易见 $\boldsymbol{\beta}-\boldsymbol{\alpha}$ 为(4)之一解，故 $\boldsymbol{\beta}-\boldsymbol{\alpha}$ 为 $\boldsymbol{\gamma}_1, \boldsymbol{\gamma}_2, \cdots, \boldsymbol{\gamma}_{n-r}$ 之线性组合：

$$\boldsymbol{\beta}-\boldsymbol{\alpha}=c_1\boldsymbol{\gamma}_1+c_2\boldsymbol{\gamma}_2+\cdots+c_{n-r}\boldsymbol{\gamma}_{n-r},$$

即

$$\boldsymbol{\beta}=\boldsymbol{\alpha}+c_1\boldsymbol{\gamma}_1+c_2\boldsymbol{\gamma}_2+\cdots+c_{n-r}\boldsymbol{\gamma}_{n-r}.$$

反之，不论 $c_i(i=1, \cdots, n-r)$ 为任何数，由于 $\boldsymbol{\alpha}$ 为(3)的解，$\sum\limits_{i=1}^{n-r}c_i\boldsymbol{\gamma}_i$ 为(5)之解，故 $\boldsymbol{\beta}=\boldsymbol{\alpha}+\sum\limits_{i=1}^{n-r}c_i\boldsymbol{\gamma}_i$ 为(3)之解．这就证得了：

定理4 当非齐次方程组(3)有解时，则(3)之任一解必为(3)之某固定解与相应齐次方程组(4)之一解的和．而且，用这样的方法可得到(3)的全部解．

练　习　八

1．试证明：设 $\alpha_1, \alpha_2, \cdots, \alpha_k$ 是某齐次线性方程组的基础解系．若 k 是奇数，则 $\alpha_1+\alpha_2, \alpha_2+\alpha_3, \cdots, \alpha_{k-1}+\alpha_k, \alpha_k+\alpha_1$ 也是该齐次线性方程组的基础解系．

2．设线性方程组

$$\begin{cases} a_{11}x_1+a_{12}x_2+\cdots+a_{1n}x_n=0, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{m1}x_1+a_{m2}x_2+\cdots+a_{mn}x_n=0 \end{cases}$$

的解都是

$$b_1x_1+b_2x_2+\cdots+b_nx_n=0$$

的解，试证 $\beta=(b_1, b_2, \cdots, b_n)$ 是

$$\alpha_1=(a_{11}, a_{12}, \cdots, a_{1n}), \ \alpha_2=(a_{21}, a_{22}, \cdots, a_{2n}), \ \cdots, \ \alpha_m=(a_{m1}, a_{m2}, \cdots, a_{mn})$$

的线性组合．

3．试证明：设 $\alpha_1=(x_1', \cdots, x_n'), \cdots, \alpha_k=(x_1^{(k)}, \cdots, x_n^{(k)})$ 为非齐次线性方程组的任 k 个解，则 $\sum\limits_{i=1}^{k} c_i\alpha_i$ 仍为这方程组之解的充要条件是 $\sum\limits_{i=1}^{k} c_i=1$．

4．解非齐次方程组：

$$\begin{cases} x_1+2x_2+x_3-x_4=4, \\ 3x_1+6x_2-x_3-3x_4=8, \\ 5x_1+10x_2+x_3-5x_4=16. \end{cases}$$

如有解，求其一组基础解系．

5．给了向量组 $x_i=(a_{i1}, a_{i2}, \cdots, a_{in})$，$(i=1, 2, \cdots, s; s\leqslant n)$．如果

$$|a_{jj}|>\sum_{\substack{i=1\\ i\neq j}}^{s}|a_{ij}|, \quad (j=1, 2, \cdots, s)$$

那末 s 个向量 x_i 一定是线性无关的．

6．证明包含 $n+2$ 个未知量 $x_1, x_2, \cdots, x_n, x_{n+1}, x_{n+2}$ 的 n 个线性齐次方程式

$$a_{i1}x_1+\cdots+a_{i,n+2}x_{n+2}=0 \ (i=1, 2, \cdots, n)$$

中总可以找到两个未知量 x_p、x_q $(p<q)$，使以任意值代入之后而所得的方程组恒有解存在．

7. 解线性方程组

$$\begin{cases}(n-1)x_1=x_2+x_3+\cdots+x_n,\\(n-1)x_2=x_1+x_3+\cdots+x_n,\\\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\(x-1)x_n=x_1+x_2+\cdots+x_{n-1}.\end{cases}$$

第4节　初 等 变 换

在第 2 节和第 3 节，我们解答了引言中提出的三个问题：判别一个线性方程组是否有解，求解的一般规则，以及讨论解之间的关系．但是，实际引用上述方法去解线性方程组时，需要计算大量的行列式，这将带来很大的困难．现在介绍简化计算的一个方法，这就是初等变换．

线性方程组

$$\begin{cases}a_{11}x_1+\cdots+a_{1n}x_n=b_1,\\\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\a_{m1}x_1+\cdots+a_{mn}x_n=b_m\end{cases}\tag{3}$$

经过下列若干个变换：(i) 变换方程的顺序，(ii) 用一个非零的数乘一方程，或(iii) 用一个数乘一方程加到另一方程上，变为

$$\begin{cases}b_{11}x_1+\cdots+b_{1n}x_n=b'_1,\\\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\b_{m1}x_1+\cdots+b_{mn}x_n=b'_m.\end{cases}\tag{10}$$

显然，所有这些变换都不影响它的解．所以，凡是(3)的解都是(10)的解；反过来，(10)又同样可以化成为(3)，因此凡是(10)的解又都是(3)的解．这就是说，(3)、(10)是同解方程组(或等价的方程组)．因此，要解(3)只须解(10)就可以了．

如果把方程组换成矩阵 $\boldsymbol{A}$，那末上面的变换就是：互换 $\boldsymbol{A}$ 的两行，用一个不为零的数乘 $\boldsymbol{A}$ 的一行，或用一个数乘 $\boldsymbol{A}$

的一行加到另一行上，这些变换叫做 $\boldsymbol{A}$ 的**行初等变换**.

于是，方程组(3)变为方程组(10)，就是(3)之增广矩阵 $\boldsymbol{A}$ 经过行初等变换变为(10)的增广矩阵 $\boldsymbol{B}$. 关于矩阵的行初等变换，有下面的一个重要性质.

定理1 *若矩阵 $\boldsymbol{A}$ 经过若干个行初等变换变为矩阵 $\boldsymbol{B}$，那末 $\boldsymbol{A}$, $\boldsymbol{B}$ 的秩相等.*

证明 这只要证明每一种行初等变换都不改变矩阵的秩就可以了. 显然，前两种行初等变换不会改变矩阵之秩. 因此，下面只就后面一种行初等变换来证明.

设用数 k 乘矩阵 $\boldsymbol{A}$ 之第 i 行加到第 j 行上，得到矩阵 $\boldsymbol{B}$，并假定 $\boldsymbol{A}$ 之秩等于 r.

容易知道：$\boldsymbol{B}$ 中任意一个 $r+1$ 级子式 B_1 若不含 $\boldsymbol{B}$ 的第 j 行，它就是 $\boldsymbol{A}$ 的一个 $r+1$ 级子式，因此 $B_1=0$; 如果 B_1 含 $\boldsymbol{B}$ 之第 j 行同时又含 $\boldsymbol{B}$ 之第 i 行，由第一章第2节中的行列式性质4得知，B_1 与 $\boldsymbol{A}$ 中一个 $r+1$ 级子式相等，所以这时也有 $B_1=0$; 如果 B_1 含 $\boldsymbol{B}$ 之第 j 行但不含 $\boldsymbol{B}$ 之第 i 行，根据行列式的性质3(参见第一章第2节)，我们有 $B_1=A_1+kA_2$，这里 A_1, A_2 都是 $\boldsymbol{A}$ 的 $r+1$ 级子式，所以也得到 $B_1=0$. 因为 B_1 在 $\boldsymbol{B}$ 中的位置只有这三种情形，所以 $\boldsymbol{B}$ 之任意 $r+1$ 级子式都是零. 这说明了 $\boldsymbol{B}$ 之秩 $\leqslant\boldsymbol{A}$ 之秩.

由于用 $-k$ 乘 $\boldsymbol{B}$ 之第 i 行加到第 j 行所得的矩阵显然就是 $\boldsymbol{A}$，于是，仿上面一段所述，又可知 $\boldsymbol{A}$ 之秩 $\leqslant\boldsymbol{B}$ 之秩.

因此，$\boldsymbol{A}$ 之秩等于 $\boldsymbol{B}$ 之秩，定理1获证.

显然，$\boldsymbol{B}$ 愈简单，(10)便愈易求解. 譬如，当 $\boldsymbol{B}$ 为最简单的对角形矩阵时，(10)的解不用计算即可看出. 然而 $\boldsymbol{B}$ 究竟能取怎样的简单形状呢?

[例1] 试用行初等变换，简化矩阵

$$A=\begin{pmatrix}1 & -2 & -1 & -2 & 2\\ 4 & 1 & 2 & 1 & 3\\ 2 & 5 & 4 & -1 & 0\\ 1 & 1 & 1 & 1 & \frac{1}{3}\end{pmatrix}.$$

解 先用 -4, -2 及 -1 乘第 1 行分别加到第 2 行、第 3 行及第 4 行上，就得到

$$\begin{pmatrix}1 & -2 & -1 & -2 & 2\\ 0 & 9 & 6 & 9 & -5\\ 0 & 9 & 6 & 3 & -4\\ 0 & 3 & 2 & 3 & -\frac{5}{3}\end{pmatrix}.$$

再用 -1 及 $-\frac{1}{3}$ 乘第 2 行分别加到第 3 行及第 4 行上，就成为

$$\begin{pmatrix}1 & -2 & -1 & -2 & 2\\ 0 & 9 & 6 & 9 & -5\\ 0 & 0 & 0 & -6 & 1\\ 0 & 0 & 0 & 0 & 0\end{pmatrix}.$$

又用 $\frac{2}{9}$ 乘第 2 行加到第 1 行上，然后用 $\frac{1}{9}$ 乘第 2 行，用 $-\frac{1}{6}$ 乘第 3 行，就有

$$\begin{pmatrix}1 & 0 & \frac{1}{3} & 0 & \frac{8}{9}\\ 0 & 1 & \frac{2}{3} & 1 & -\frac{5}{9}\\ 0 & 0 & 0 & 1 & -\frac{1}{6}\\ 0 & 0 & 0 & 0 & 0\end{pmatrix}.$$

最后，用 -1 乘第 3 行加到第 2 行，得

$$\begin{pmatrix} 1 & 0 & \frac{1}{3} & 0 & \frac{8}{9} \\ 0 & 1 & \frac{2}{3} & 0 & -\frac{7}{18} \\ 0 & 0 & 0 & 1 & -\frac{1}{6} \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

这就是所给矩阵的简单形状，它的秩等于 3；它的特点是：每个列向量的第 3 个分量以后的分量都是零，并且有三列只有一个分量为 1，其他分量都是零，这三个列是

$$\begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \quad \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}.$$

对于一般情况，因为矩阵 $\boldsymbol{A}$ 的任意非零列向量都可以用行初等变换化为只有一个分量是 1，其他分量都是零的列向量，所以假如 $\boldsymbol{A}$ 之秩等于 r，我们可以用行初等变换把它的某 r 个列向量化为只有 1 个分量是 1，其他各分量都是零的列向量，并且所有列向量的第 r 个分量的后面的分量都是零．

定理 2 满秩矩阵恒可用行初等变换化为单位矩阵[†]．

于是，用行初等变换把 $\boldsymbol{B}$ 化为上述形状后，(10) 只要再经过适当的移项，就可以求得它的解，这就不必计算行列式了．

[例 2] 解线性方程组

---

† 主对角线上都是 1，其他位置都是 0 的 n 级矩阵，叫做 n 级单位矩阵．

$$\begin{cases} x_1-2x_2-x_3-2x_4=2, \\ 4x_1+x_2+2x_3+x_4=3, \\ 2x_1+5x_2+4x_3-x_4=0, \\ x_1+x_2+x_3+x_4=\dfrac{1}{3}. \end{cases}$$

解 方程组的增广矩阵正是例 1 中矩阵，因此，用行初等变换可以化成为

$$\begin{pmatrix} 1 & 0 & \dfrac{1}{3} & 0 & \dfrac{8}{9} \\ 0 & 1 & \dfrac{2}{3} & 0 & -\dfrac{7}{18} \\ 0 & 0 & 0 & 1 & -\dfrac{1}{6} \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

于是所给线性方程组就化为

$$\begin{cases} x_1+\dfrac{1}{3}x_3=\dfrac{8}{9}, \\ x_2+\dfrac{2}{3}x_3=-\dfrac{7}{18}, \\ x_4=-\dfrac{1}{6}, \end{cases}$$

移项，就得到

$$x_1=\frac{8}{9}-\frac{1}{3}x_3,\ x_2=-\frac{7}{18}-\frac{2}{3}x_3,\ x_4=-\frac{1}{6}$$

(x_3 得任意取值)，

这就是所求之解.

注意：用行初等变换解线性方程组时，不必先计算系数矩阵及增广矩阵的秩，在变换过程中如发现两个矩阵的秩不等，那末方程组就无解.

[例 3] 解线性方程组

$$\begin{cases} x+2y-z=1, \\ 2x-3y+z=0, \\ 4x+y-z=-1. \end{cases}$$

解

$$\boldsymbol{B}=\begin{pmatrix} 1 & 2 & -1 & 1 \\ 2 & -3 & 1 & 0 \\ 4 & 1 & -1 & -1 \end{pmatrix}.$$

先用 -2 和 -4 乘第 1 行分别加到第 2 和第 3 行上，得

$$\begin{pmatrix} 1 & 2 & -1 & 1 \\ 0 & -7 & 3 & -2 \\ 0 & -7 & 3 & -5 \end{pmatrix}.$$

再用 -1 乘第 2 行加到第 3 行，得

$$\begin{pmatrix} 1 & 2 & -1 & 1 \\ 0 & -7 & 3 & -2 \\ 0 & 0 & 0 & -3 \end{pmatrix}.$$

由上可见：方程组之系数矩阵的秩与增广矩阵的秩不相等，所以方程组无解.

一般，对一个矩阵，除行初等变换外，同样还有*列初等变换*，我们把行初等变换及列初等变换统称为初等变换，即

定义 矩阵 $\boldsymbol{A}$ 的下列变换，叫数 $\boldsymbol{A}$ 的*初等变换*：

1° 互换 $\boldsymbol{A}$ 之两行或两列；

2° 用一个不为 0 的数乘 $\boldsymbol{A}$ 的一行或一列；

3° 用一个数乘 $\boldsymbol{A}$ 的一行加到另一行上，或乘一列加到另一列上.

若矩阵 $\boldsymbol{A}$ 经过有穷个初等变换化为矩阵 $\boldsymbol{B}$，就叫 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 等价，用 $\boldsymbol{A}\simeq\boldsymbol{B}$ 表示. 由定义，容易得知，对于“等价”这个

关系，下面三个定律成立：

i. 自反律：$A\simeq A$；

ii. 对称律：若 $A\simeq B$，则 $B\simeq A$；

iii. 传递律：若 $A\simeq B$，$B\simeq C$，则 $A\simeq C$.

一般，满足这三个律的关系叫做*等价关系*. 例如，初等几何中三角形相似的性质是三角形的等价关系.

前面，用行初等变换可把矩阵简化，如再用列初等变换，可使矩阵再简化. 譬如例 1 中矩阵，再用列初等变换，用 $-\frac{1}{3}$ 乘第 1 列、又用 $-\frac{2}{3}$ 乘第 2 列，都加到第 3 列上，再用 $-\frac{8}{9}$ 乘第 1 列、用 $\frac{7}{18}$ 乘第 2 列、用 $\frac{1}{6}$ 乘第 4 列，都加到第 5 列上，最后把第 3 列与第 4 列互换，就得到

$$\begin{pmatrix}1 & 0 & \frac{1}{3} & 0 & \frac{8}{9}\\ 0 & 1 & \frac{2}{3} & 0 & -\frac{7}{18}\\ 0 & 0 & 0 & 1 & -\frac{1}{6}\\ 0 & 0 & 0 & 0 & 0\end{pmatrix}\simeq\begin{pmatrix}1 & 0 & 0 & 0 & 0\\ 0 & 1 & 0 & 0 & 0\\ 0 & 0 & 1 & 0 & 0\\ 0 & 0 & 0 & 0 & 0\end{pmatrix}.$$

这样我们就得到最简单的矩阵了.

上面的矩阵中，只有前面某些第 i 行、第 i 列上的数是 1，其余都是零. 这种简单形状的矩阵，叫做矩阵的*标准形*.

由矩阵之标准形的意义，可知：任一矩阵总可以经过行及列的若干次初等变换化为标准形. 由定理 1，经过行初等变换后，矩阵的秩不变，显然，同样可知经过列初等变换也不改变矩阵的秩. 这就是说，经过初等变换，不改变矩阵的秩. 于是，从标准形来求矩阵的秩是非常方便. 当然，求秩时不必把

矩阵化为标准形，只要适当地使用初等变换，能够看出它的秩就可以了. 这样，不需要计算行列式就可以求出矩阵的秩了.

注意：在求矩阵之秩时，用的初等变换可以为行初等变换，也可以为列初等变换，但解线性方程组只能用行初等变换而不能用列初等变换. 因为，用列初等变换就把未知量搞乱了，当然就不便求出所给线性方程组的解.

对于两个矩阵，如果它们的行数、列数分别都相等，叫做同型的. 譬如$\begin{pmatrix}1 & 2 & -1\\ 2 & -3 & 1\end{pmatrix}$与$\begin{pmatrix}0 & 2 & -4\\ -1 & -4 & 5\end{pmatrix}$同型，但不与$\begin{pmatrix}3 & -2\\ 1 & 4\end{pmatrix}$同型.

由于经过初等变换不改变矩阵的秩，故知：若两个矩阵等价，则它们的秩相等. 反之，若两个同型矩阵的秩相等，那末它们就有相同的标准形，即它们都与同一个矩阵等价，因之由传递律知它们互为等价的. 这就证得了：

定理 3 两个同型矩阵等价的充要条件是它们的秩相等.

练 习 九

1. 试用初等变换，解线性方程组：

$$\begin{cases} ax+y+z=p, \\ x+ay+z=q, \\ x+y+az=r. \end{cases}$$

2. 设 $m\times n$ 矩阵 A 经过行初等变换化为矩阵 B，A 的列向量分别用 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 表示，B 的列向量分别用 $\beta_1, \beta_2, \cdots, \beta_n$ 表示. 试证：如果

$$k_1\alpha_1+k_2\alpha_2+\cdots+k_n\alpha_n=0, \tag{*}$$

那末

$$k_1\beta_1+k_2\beta_2+\cdots+k_n\beta_n=0; \tag{**}$$

反之，若(**)成立，则(*)也成立.

3. 试证：第 1° 种初等变换可以由其他两种初等变换导出.

第三章　向 量 空 间

本章将介绍 n 维向量空间的基本概念和它的基本性质，先从直觉空间(二维、三维)的推广入手，然后再把 n 维向量空间的概念抽象化(即公理化)．这样，使读者容易接受，并且也符合发展程序．

第1节　n 维空间的意义

平常我们说二维空间，一般是指平面上所有点的集合．如果在这个平面上取定了笛卡儿坐标系之后，则平面上的点和实数对(即两个有次序的实数组)成一一对应关系，这就是说：平面上有一点，则必有维一对的有次序的实数组 (x, y) (即点的坐标)与它对应；反之，如果给了两个有次序的实数组 (x, y)，则在平面上也就有唯一的点与它相对应．而之，二维空间是一切有次序的实数对之集合．另一方面，任一个实数对 (x, y) 又可看成是平面上起点是坐标原点、终点是 (x, y) 的向量，因之二维空间也叫做二维向量空间，并满足：平面上任意两个过原点的向量之和以及数与向量之积，仍是过原点的向量．同样的道理，可以把三维空间(即空间中一切点的集合)解释为三维向量空间，在它里面向量可以相加并且也可以和数相乘．在第二章，我们已经介绍了 n 元向量的加法以及它与数的乘法，把上面的概念推广，就得到：

定义　在 n 元向量的集合 $\mathfrak{R}$ 中，如果其中任意两个向量

的和以及任意一个向量与数的积都仍在 $\mathfrak{R}$ 中，那末 $\mathfrak{R}$ 叫做向量空间.

显然，所有 n 元向量的集合形成一个向量空间，通常用符号 $\mathfrak{R}_n$ 来表示. 只包含一个零向量的集合也是一个向量空间，叫做零空间. n 个未知量的线性齐次方程组的解向量的全体也形成一个向量空间，叫做这方程组的解空间. 不过，线性非齐次方程组的解向量的全体却不能形成一个向量空间.

注意：n 元向量 $(x_1, x_2, \cdots, x_n)$ 的分量 x_i 可以是任意的数. 当 x_i 都是实数时，这样一些向量形成的向量空间叫做实空间，否则就叫做复空间(即 x_i 允许有为复数的可能). 通常说的向量空间，指的都是复空间，而叫 $\mathfrak{R}_n$(即一切 n 元复向量 $(x_1, x_2, \cdots, x_n)$ 之集合)为 n 维(复)向量空间；若特限制每个 x_i 为实数，即一切 n 元实向量 $(x_1, \cdots, x_n)$ 之集合虽仍用 $\mathfrak{R}_n$ 表示，但这时就特地叫它为 n 维实空间，或简称实 $\mathfrak{R}_n$. 又，有时候常将“向量”二字省掉，也有只说空间(指复空间)或实空间的，这时 n 个数组 $(x_1, x_2, \cdots, x_n)$ 通常又叫做空间的点. 故今后说 n 个数组 $(x_1, x_2, \cdots, x_n)$ 为空间中的点也可以，或者说它是空间中的向量也可以，看实际的需要来决定. 不论叫点或叫向量，它都没有直觉的几何意义(当 $n\geq 4$ 时)，而仅是将直觉空间中($n=1, 2, 3$ 时)一些几何问题变成代数关系式，再将这些代数关系式推广到 $n>3$ 的情形而沿用几何的术语而已. 例如，空间中有两个点 $A(x_1, x_2, x_3)$, $B(y_1, y_2, y_3)$, 点 $C(z_1, z_2, z_3)$ 是直线 AB 上的任意一点(图 3-1)，于是 $\overrightarrow{AC}$ 显然等于 $\lambda\cdot\overrightarrow{AB}$ (实数 λ 是随 C 点之位置的变化而变的). 于是

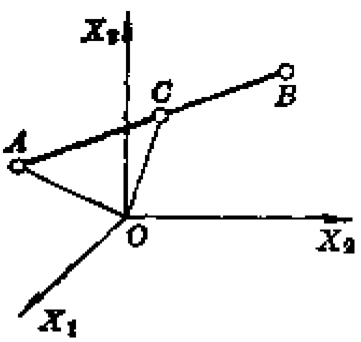


图 3-1

$$\overrightarrow{OC}=\overrightarrow{OA}+\overrightarrow{AC}=\overrightarrow{OA}+\lambda\cdot\overrightarrow{AB}.$$

但 $$\overrightarrow{OC}=\{z_1,\ z_2,\ z_3\},\quad \overrightarrow{OA}=\{x_1,\ x_2,\ x_3\},$$
$$\overrightarrow{AB}=\{y_1-x_1,\ y_2-x_2,\ y_3-x_3\},$$

故有

$$\{z_1, z_2, z_3\}=\{x_1+\lambda(y_1-x_1), x_2+\lambda(y_2-x_2), x_3+\lambda(y_3-x_3)\},$$

所以

$$\begin{cases} z_1=x_1+\lambda(y_1-x_1), \\ z_2=x_2+\lambda(y_2-x_2), \\ z_3=x_3+\lambda(y_3-x_3). \end{cases} \tag{1}$$

在(1)式中，若使 λ 取一切实数值时，可得过 A, B 之直线上一切点的坐标；若使 λ 取一切位于 0 至 1 间的实数($0\leqslant\lambda\leqslant1$)，则可得线段 AB 上的一切点之坐标．我们叫(1)式为**直线** AB **的参数方程**．这都是大家所熟悉的解析几何的内容．现在将这概念推广到 n 维空间 $\mathfrak{R}_n$:

设 $P=(x_1,\ x_2,\ \cdots,\ x_n)$, $Q=(y_1,\ y_2,\ \cdots,\ y_n)$ 为空间 $\mathfrak{R}_n$ 中两点，那末所有适合方程

$$z_i=x_i+\lambda(y_i-x_i)\quad (i=1,\ 2,\ \cdots,\ n) \tag{2}$$

(λ 为任何实数)的点$(z_1,\ z_2,\ \cdots,\ z_n)$之集合叫做 $\mathfrak{R}_n$ 中**通过** P, Q **两点的直线**；如果 $0\leqslant\lambda\leqslant1$，则凡适合上列方程(2)的点之集合叫做**线段** PQ.

练　习　十

1．在实空间 $\mathfrak{R}_n$ 中设有两点 $P(x_1,\ x_2,\ \cdots,\ x_n)$ 与 $Q(y_1,\ y_2,\ \cdots,\ y_n)$，而 R 与 S 是线段 PQ 上的某两点，证明线段 RS 是线段 PQ 的一部分，即线段 RS 上的点也一定在线段 PQ 上．

2．在上题中，证明通过 R, S 两点的直线与通过 P, Q 两点的直线是一致的．

第2节　基底、维数及坐标

大家知道，在平面直角坐标系中，OX 轴和 OY 轴上分别取坐标单位向量 $\boldsymbol{i}$ 和 $\boldsymbol{j}$(即它们的长等于1：$|\boldsymbol{i}|=1$，$|\boldsymbol{j}|=1$；它们的方向分别和所在的轴的正向相同)，如图 3-2，于是对于平面上的任意向量 $\boldsymbol{a}$ $(=\overrightarrow{OA})$ 显然可以写为

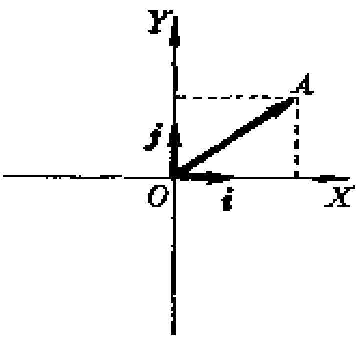


图　3-2

$$\boldsymbol{a}=x\boldsymbol{i}+y\boldsymbol{j},$$

并且，这样表示的方法还是唯一的．这就是向量 $\boldsymbol{a}$ 按坐标单位向量的分解．式中的 x 和 y，实际上是向量 $\boldsymbol{a}$ 的终点 A 的坐标．我们叫向量 $\boldsymbol{i}$、$\boldsymbol{j}$ 为这向量空间(二维)的基底，表示式 $\boldsymbol{a}=x\boldsymbol{i}+y\boldsymbol{j}$ 为向量 $\boldsymbol{a}$ 按基底 $\boldsymbol{i}$, $\boldsymbol{j}$ 的分解；向量 $x\boldsymbol{i}$, $y\boldsymbol{j}$ 叫做向量 $\boldsymbol{a}$ 在基底 $\boldsymbol{i}$, $\boldsymbol{j}$ 上的分向量．当然，两向量 $\boldsymbol{i}$, $\boldsymbol{j}$ 是线性无关的(因为它们不共线)，若用坐标来表示，就是 $\boldsymbol{i}=\{1, 0\}$，$\boldsymbol{j}=\{0, 1\}$.

同样的道理，如图 3-3，空间的每个向量 $\boldsymbol{a}(=\overrightarrow{OA})$ 可以唯一地写成

$$\boldsymbol{a}=x\boldsymbol{i}+y\boldsymbol{j}+z\boldsymbol{k}$$

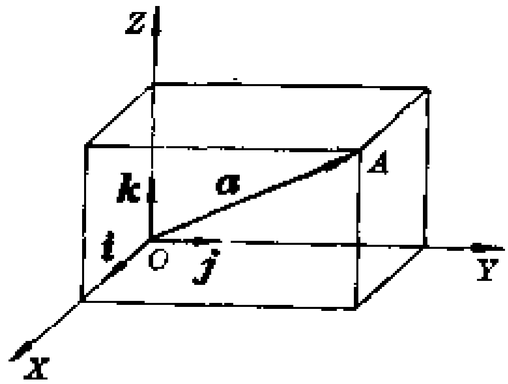


图　3-3

的形式，这里 x, y, z 是向量 $\boldsymbol{a}=\overrightarrow{OA}$ 的终点 A 的坐标．我们叫向量 $\boldsymbol{i}$, $\boldsymbol{j}$, $\boldsymbol{k}$ 是这空间的基底，表示式 $\boldsymbol{a}=x\boldsymbol{i}+y\boldsymbol{j}+z\boldsymbol{k}$ 叫做向量 $\boldsymbol{a}$ 按基底 $\boldsymbol{i}$, $\boldsymbol{j}$, $\boldsymbol{k}$ 的分解；向量 $x\boldsymbol{i}$, $y\boldsymbol{j}$, $z\boldsymbol{k}$ 叫做向量 $\boldsymbol{a}$ 在基底 $\boldsymbol{i}$, $\boldsymbol{j}$, $\boldsymbol{k}$ 上的分向量．当然，$\boldsymbol{i}$, $\boldsymbol{j}$, $\boldsymbol{k}$ 是线性无关的(因为它们不共面)，用坐标来表示，就有 $\boldsymbol{i}=\{1, 0, 0\}$，$\boldsymbol{j}=\{0, 1,$

0}, $\boldsymbol{k}=\{0, 0, 1\}$.

上面说的都是直觉空间(二维或三维)中的事．然 $n>3$ 时，非直觉的 n 维向量空间 $\mathfrak{R}_n$ 中也存在 n 个线性无关的向量：

$$\boldsymbol{e}_1=(1, 0, \cdots, 0), \boldsymbol{e}_2=(0, 1, 0, \cdots, 0), \cdots\cdots, \boldsymbol{e}_n=(0, \cdots, 0, 1),$$

使得 $\mathfrak{R}_n$ 中每个(n 元)向量 $\boldsymbol{\alpha}=(a_1, a_2, \cdots, a_n)$都可以唯一地写成 $\boldsymbol{e}_1, \boldsymbol{e}_2, \cdots, \boldsymbol{e}_n$ 的线性组合，即

$$\boldsymbol{\alpha}=a_1\boldsymbol{e}_1+a_2\boldsymbol{e}_2+\cdots+a_n\boldsymbol{e}_n.$$

这说明 $\boldsymbol{e}_1, \boldsymbol{e}_2, \cdots, \boldsymbol{e}_n$ 在 $\mathfrak{R}_n$ 内的作用恰与 $\boldsymbol{i}, \boldsymbol{j}$ 在二维空间或 $\boldsymbol{i}, \boldsymbol{j}, \boldsymbol{k}$ 在三维空间内的作用完全一样．因此，n 个向量 $\boldsymbol{e}_1, \boldsymbol{e}_2, \cdots, \boldsymbol{e}_n$ 就是 $\mathfrak{R}_n$ 的基底．下面更一般地介绍基底这个概念，这是一个重要的概念，它在向量空间理论中的作用，正象解析几何里的平面直角坐标系．

定义 设 $\mathfrak{R}$ 是由 n 元向量形成的向量空间，$\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_m$ 是向量空间 $\mathfrak{R}$ 的 m 个线性无关的向量，如果 $\mathfrak{R}$ 的每个向量 $\boldsymbol{\alpha}$ 都可以写成它们的线性组合，如 $\boldsymbol{\alpha}=c_1\boldsymbol{\alpha}_1+\cdots+c_m\boldsymbol{\alpha}_m$，我们就叫 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_m$ 是 $\mathfrak{R}$ 的*基底*(或*一个基底*)．

由这定义，就会提出下面四个问题．

1) $\mathfrak{R}$ 的基底是否存在？这个问题的答案是肯定的．因为 $\mathfrak{R}$ 是 $\mathfrak{R}_n$ 之子集合，而 $\mathfrak{R}_n$ 之每向量(n 元的)均可为前述的 $\boldsymbol{e}_1, \boldsymbol{e}_2, \cdots, \boldsymbol{e}_n$ 的线性组合，故 $\mathfrak{R}_n$ 随而 $\mathfrak{R}$ 中的任 $n+1$ 个(n 元)向量都是线性相关的，因而在 $\mathfrak{R}$ 内必可找着 $m(\leqslant n)$个线性无关的向量，使 $\mathfrak{R}$ 中任意向量都为这 m 个向量的线性组合，这就说明了 $\mathfrak{R}$ 之基底确实存在．

根据前章第1节，易知上述之 m 为 $\mathfrak{R}$ 中的一个定数，我们叫这个定数 m 为 $\mathfrak{R}$ 的*维数*．于是，由 n 元向量形成的向量

空间之维数不大于 n，而维数等于 n 之充要条件是 $\mathfrak{R}=\mathfrak{R}_n$. 又，零空间没有基底，我们说它的维数是零. 假如 n 个未知量的齐次线性方程组的秩是 r，那末它的基础解系所含解向量的个数为 $n-r$，因此解空间的维数等于 $n-r$.

2) 在定义中，$\boldsymbol{\alpha}=c_1\boldsymbol{\alpha}_1+\cdots+c_m\boldsymbol{\alpha}_m$ 的表示方法是唯一的吗？这答案也是肯定的，因为从

$$\boldsymbol{\alpha}=c_1\boldsymbol{\alpha}_1+\cdots+c_m\boldsymbol{\alpha}_m=d_1\boldsymbol{\alpha}_1+\cdots+d_m\boldsymbol{\alpha}_m,$$

即得
$$(c_1-d_1)\boldsymbol{\alpha}_1+\cdots+(c_m-d_m)\boldsymbol{\alpha}_m=\mathbf{0},$$

故据 $\boldsymbol{\alpha}_1, \cdots, \boldsymbol{\alpha}_m$ 线性无关，得 $c_i-d_i=0$ 即 $c_i=d_i(i=1, \cdots, m)$. 正因为这种唯一性，我们把线性表示的系数 $c_1, \cdots, c_m$ 叫做 $\boldsymbol{\alpha}$ 关于基底 $\boldsymbol{\alpha}_1, \cdots, \boldsymbol{\alpha}_m$ 的坐标.

注意：对于不同的基底，同一个向量的坐标是不同的. 在 $\mathfrak{R}_n$ 中，假如取 $\boldsymbol{\alpha}_1=(1, 1, \cdots, 1)$，$\boldsymbol{\alpha}_2=(0, 1, \cdots, 1), \cdots, \boldsymbol{\alpha}_n=(0, \cdots, 0, 1)$做基底，因为

$$\begin{aligned}\boldsymbol{\alpha}&=(a_1, a_2, \cdots, a_n)\\&=a_1\boldsymbol{\alpha}_1+(a_2-a_1)\boldsymbol{\alpha}_2+\cdots+(a_n-a_{n-1})\boldsymbol{\alpha}_n,\end{aligned}$$

故 $\boldsymbol{\alpha}$ 关于基底 $\boldsymbol{\alpha}_1, \cdots, \boldsymbol{\alpha}_n$ 的坐标就是 $a_1, a_2-a_1, \cdots, a_n-a_{n-1}$. 如果取 $\boldsymbol{e}_1, \boldsymbol{e}_2, \cdots, \boldsymbol{e}_n$ 做基底，因为

$$\boldsymbol{\alpha}=(a_1, a_2, \cdots, a_n)=a_1\boldsymbol{e}_1+\cdots+a_n\boldsymbol{e}_n,$$

故 $\boldsymbol{\alpha}$ 关于 $\boldsymbol{e}_1, \cdots, \boldsymbol{e}_n$ 的坐标就是 $a_1, a_2, \cdots, a_n$.

据上述得知：由 n 元向量之集合所组成的向量空间的维数 $m\leqslant n$，$m<n$ 的可能性是有的.

注意：所谓 m 维向量空间，其中的向量都是 n 元向量，并非 m 元向量. 譬如，零空间的维数是零，它所含的唯一的零向量是 n 元向量. m 维向量空间的基底所含向量的个数是 m，因此对于它的基底 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_m$（每个 $\boldsymbol{\alpha}_i$ 为 n 元向量），可以把 m 维空间中的 n 元向量 $\boldsymbol{\alpha}=(a_1, a_2, \cdots, a_n)$ 表示成

$\boldsymbol{\alpha}=c_1\boldsymbol{\alpha}_1+\cdots+c_m\boldsymbol{\alpha}_m$，因而又可记为 $\boldsymbol{\alpha}=(c_1, c_2, \cdots, c_m)$. 虽然，我们应特别注意 n 元向量 $\boldsymbol{\alpha}=(a_1, a_2, \cdots, a_n)$ 写成 m 元向量 $(c_1, c_2, \cdots, c_m)$ 之含义.

3) 如果已知 $\mathfrak{R}_n$ 之维数等于 n, 但 $\mathfrak{R}_n$ 之基底可以不止一组，试问 $\mathfrak{R}_n$ 中 n 个向量 $\boldsymbol{\alpha}_i=(a_{i1}, a_{i2}, \cdots, a_{in})$, $(i=1, 2, \cdots, n)$, 在什么条件下才能够是 $\mathfrak{R}_n$ 之基底呢?

易知它们为 $\mathfrak{R}_n$ 之基底的充要条件是它们为线性无关的，因而 $\boldsymbol{\alpha}_i$(其中 $i=1, \cdots, n$) 为基底的充要条件是由其坐标 a_{i1}, a_{i2}, $\cdots$, a_{in} 组成的行列式不等于零，于是有下面的两个定理.

定理 1 n 维空间 $\mathfrak{R}_n$ 中 n 个向量 $\boldsymbol{\alpha}_i=(a_{i1}, a_{i2}, \cdots, a_{in})$, $(i=1, 2, \cdots, n)$, 是 $\mathfrak{R}_n$ 之一组基底的充要条件是

$$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} \neq 0.$$

同理，也可以证明

定理 2 设在由 n 元向量形成的 m 维空间 $\mathfrak{R}$ 中取一组基底 $\boldsymbol{\alpha}_1, \cdots, \boldsymbol{\alpha}_m$(都是 n 元向量)，那末 $\mathfrak{R}$ 里面 m 个向量 $\boldsymbol{\beta}_1, \cdots, \boldsymbol{\beta}_m$ (都是 n 元向量) 是 $\mathfrak{R}$ 之基底的充要条件是它们关于 $\boldsymbol{\alpha}_1, \cdots, \boldsymbol{\alpha}_m$ 的坐标所形成的 m 级行列式不等于零，即当 $\boldsymbol{\beta}_i=c_{i1}\boldsymbol{\alpha}_1+\cdots+c_{im}\boldsymbol{\alpha}_m$(其中 $i=1, \cdots, m$) 时，

$$\begin{vmatrix} c_{11} & c_{12} & \cdots & c_{1m} \\ c_{21} & c_{22} & \cdots & c_{2m} \\ \vdots & \vdots & & \vdots \\ c_{m1} & c_{m2} & \cdots & c_{mm} \end{vmatrix} \neq 0.$$

4) n 维空间 $\mathfrak{R}_n$ 中任意 k 个 $(k<n)$ 线性无关的向量 $\boldsymbol{\beta}_1, \cdots, \boldsymbol{\beta}_k$ 能够是 $\mathfrak{R}_n$ 之基底的一部分吗? 也就是说，能够在

$\boldsymbol{\beta}_1, \cdots, \boldsymbol{\beta}_k$ 上再添加 $n-k$ 个线性无关的向量，使它们成为 $\mathfrak{R}_n$ 之基底吗？这也是肯定的．事实上，由 $\boldsymbol{\beta}_i=(b_{i1}, b_{i2}, \cdots, b_{in})$, $(i=1, 2, \cdots, k)$，可形成一个有 k 行 n 列的矩阵

$$\boldsymbol{B}=\begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & & \vdots \\ b_{k1} & b_{k2} & \cdots & b_{kn} \end{pmatrix},$$

其秩等于 k，不失普遍性，可令

$$\begin{vmatrix} b_{11} & \cdots & b_{1k} \\ \vdots & & \vdots \\ b_{k1} & \cdots & b_{kk} \end{vmatrix} \neq 0,$$

于是据拉普拉斯展开式，得知

$$\begin{vmatrix} b_{11} & \cdots & b_{1k} & b_{1,k+1} & \cdots & b_{1n} \\ \vdots & & \vdots & \vdots & & \vdots \\ b_{k1} & \cdots & b_{kk} & b_{k,k+1} & \cdots & b_{kn} \\ 0 & \cdots & 0 & 1 & \cdots & 0 \\ \vdots & & \vdots & \vdots & & \vdots \\ 0 & \cdots & 0 & 0 & \cdots & 1 \end{vmatrix} \neq 0,$$

即 $\boldsymbol{\beta}_1, \cdots, \boldsymbol{\beta}_k, \boldsymbol{e}_{k+1}, \cdots, \boldsymbol{e}_n$ 可为 $\mathfrak{R}_n$ 之一组基底．

同样，对 n 元向量所形成的 m 维空间 $\mathfrak{R}$，也有如下类似的结果．

定理 8 由 n 元向量所形成的 m 维空间 $\mathfrak{R}(m\leqslant n;\ m=n$ 时，$\mathfrak{R}=\mathfrak{R}_n)$ 中任意 k 个线性无关的向量（n 元向量）$(k\leqslant m)$ 是 $\mathfrak{R}$ 之基底的一部分．

证明 设 $\boldsymbol{\alpha}_i(i=1, 2, \cdots, k)$ 是 m 维空间 $\mathfrak{R}$ 中 k 个线性无关的向量（n 元的），并设 $\boldsymbol{\beta}_i$（其中 $i=1, 2, \cdots, m$）为 $\mathfrak{R}$ 之一组基．于是就有

$$\begin{aligned}\boldsymbol{\alpha}_1&=c_{11}\boldsymbol{\beta}_1+c_{12}\boldsymbol{\beta}_2+\cdots+c_{1m}\boldsymbol{\beta}_m,\\ \boldsymbol{\alpha}_2&=c_{21}\boldsymbol{\beta}_1+c_{22}\boldsymbol{\beta}_2+\cdots+c_{2m}\boldsymbol{\beta}_m,\\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\ \boldsymbol{\alpha}_k&=c_{k1}\boldsymbol{\beta}_1+c_{k2}\boldsymbol{\beta}_2+\cdots+c_{km}\boldsymbol{\beta}_m,\end{aligned}$$

且矩阵$\begin{pmatrix} c_{11} & c_{12} & \cdots & c_{1m} \\ c_{21} & c_{22} & \cdots & c_{2m} \\ \vdots & \vdots & & \vdots \\ c_{k1} & c_{k2} & \cdots & c_{km}\end{pmatrix}$之行向量线性无关，因而其秩等于$k$. 不失普遍性，可令

$$\begin{vmatrix} c_{11} & \cdots & c_{1k} \\ \vdots & & \vdots \\ c_{k1} & \cdots & c_{kk}\end{vmatrix}\neq 0,$$

随而再据拉普拉斯展开式，得知

$$\begin{vmatrix} c_{11} & \cdots & c_{1k} & c_{1,k+1} & \cdots & c_{1m} \\ \vdots & & \vdots & \vdots & & \vdots \\ c_{k1} & \cdots & c_{kk} & c_{k,k+1} & \cdots & c_{km} \\ 0 & \cdots & 0 & 1 & \cdots & 0 \\ \vdots & & \vdots & \vdots & & \vdots \\ 0 & \cdots & 0 & 0 & \cdots & 1 \end{vmatrix}\neq 0,$$

因之$\boldsymbol{\alpha}_1, \cdots, \boldsymbol{\alpha}_k, \boldsymbol{\gamma}_1(=\boldsymbol{\beta}_{k+1}), \cdots, \boldsymbol{\gamma}_{m-k}(=\boldsymbol{\beta}_m)$线性无关而为$\mathfrak{R}$之一组基底.

练 习 十 一

1. 假定$\boldsymbol{\alpha}_1=(1, -1, 2, 3)$、$\boldsymbol{\alpha}_2=(3, 0, 4, -2)$，试求$\boldsymbol{\alpha}_3$、$\boldsymbol{\alpha}_4$，使$\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_3, \boldsymbol{\alpha}_4$构成$\mathfrak{R}_4$之基底.

2. 设$\boldsymbol{\alpha}_1=(1, 1, -1, -1, 2)$，$\boldsymbol{\alpha}_2=(2, 1, 0, 1, 0)$，$\boldsymbol{\alpha}_3=(3, 2, -1, 0, 2)$. 证明：$\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_3$的一切线性组合$c_1\boldsymbol{\alpha}_1+c_2\boldsymbol{\alpha}_2+c_3\boldsymbol{\alpha}_3$($c_1, c_2, c_3$为任意数)的集合是一个向量空间．并求这空间的维数.

第3节　子　空　间

如图 3-4，在二维空间 $\mathfrak{R}_2$(即某个平面)内取定平面直角坐标系之后，引任一个不等于零之向量 $\boldsymbol{a}=\overrightarrow{OA}$，于是 $\boldsymbol{a}$ 与过原点的任一个和 $\boldsymbol{a}$ 不共线的向量之一切线性组合(实系数)的集合就是二维空间 $\mathfrak{R}_2$. 但 $\boldsymbol{a}$ 之一切倍数 $\lambda\boldsymbol{a}$(λ 是任意的实数)的集合显然是通过 O, A 两点的直线 l. 根据向量空间及其维数之定义，则知直线 l 是一维空间. 因直线 l 在平面 $\mathfrak{R}_2$ 上，所以我们就说一维空间(即直线 l)是 $\mathfrak{R}_2$ 的子空间.

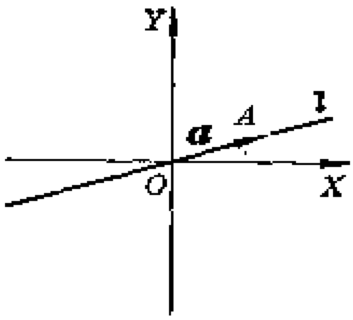


图　3-4

又如图 3-5，在三维空间 $\mathfrak{R}_3$ 内取定平面直角坐标系后，引两个不共线的向量 $\boldsymbol{a}=\overrightarrow{OA}$ 与 $\boldsymbol{b}=\overrightarrow{OB}$，则 $\boldsymbol{a}$ 与 $\boldsymbol{b}$ 的一切线性组合 $\lambda\boldsymbol{a}+\mu\boldsymbol{b}$ 的集合(λ, μ 为实数)显然是通过 O, A, B 三点的平面 π. 根据向量空间的定义，则知平面 π 为一个二维空间. 但因平面 π 完全在 $\mathfrak{R}_3$ 内，所以 π 是 $\mathfrak{R}_3$ 的子空间. 又，$\boldsymbol{a}$ 之一切倍数 $\lambda\boldsymbol{a}$ 之集合显然是通过 O, A 两点的直线 l，所以 l 是一个一维空间，又因它完全在 $\mathfrak{R}_3$ 内，所以 l 也是 $\mathfrak{R}_3$ 的子空间.

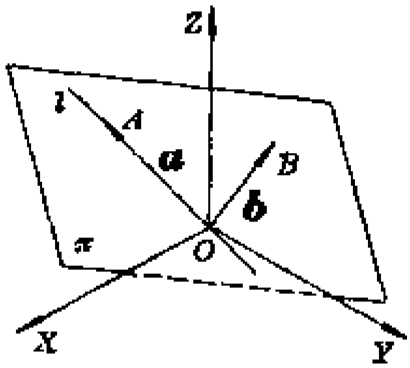


图　3-5

一般，我们有：

定义　假定 $\mathfrak{R}$ 是向量空间，$\mathfrak{S}$ 是 $\mathfrak{R}$ 中某些向量的集合. 如果 $\mathfrak{S}$ 又是向量空间，那末叫 $\mathfrak{S}$ 为 $\mathfrak{R}$ 的子空间.

对于任意向量空间，有零空间做它的子空间，向量空间本身也是它自己的子空间。又，n个未知量的线性齐次方程组的解空间，是所有n元向量形成的n维向量空间$\mathfrak{R}_n$之子空间。由向量空间的定义，我们得到

定理1 向量空间$\mathfrak{R}$中的向量的集合$\mathfrak{S}$，是$\mathfrak{R}$的子空间的充要条件是：

i) 如果$\boldsymbol{\alpha}\in\mathfrak{S}$，那末$\lambda\boldsymbol{\alpha}\in\mathfrak{S}$($\lambda$是任意数)；

ii) 如果$\boldsymbol{\alpha}$, $\boldsymbol{\beta}\in\mathfrak{S}$，那末$\boldsymbol{\alpha}+\boldsymbol{\beta}\in\mathfrak{S}$。

下面来谈子空间的形成。

假定$\boldsymbol{\beta}_1, \cdots, \boldsymbol{\beta}_k$是向量空间$\mathfrak{R}$中的向量。显然，形如

$$c_1\boldsymbol{\beta}_1+c_2\boldsymbol{\beta}_2+\cdots+c_k\boldsymbol{\beta}_k$$

的向量的和，以及它与数的积，都还是这种形状的向量。因此，所有这种向量的集合$\mathfrak{S}$，形成一个向量空间。于是，我们得到

定理2 设$\boldsymbol{\beta}_1, \cdots, \boldsymbol{\beta}_k$是向量空间$\mathfrak{R}$中的向量，那末$\boldsymbol{\beta}_1, \cdots, \boldsymbol{\beta}_k$的一切线性组合$c_1\boldsymbol{\beta}_1+\cdots+c_k\boldsymbol{\beta}_k$的集合$\mathfrak{S}$形成$\mathfrak{R}$的一个子空间。

定理2中的$\mathfrak{S}$，称为由$\boldsymbol{\beta}_1, \cdots, \boldsymbol{\beta}_k$生成的子空间，用符号$\mathfrak{S}=\mathfrak{S}[\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \cdots, \boldsymbol{\beta}_k]$记之。

再探讨$\mathfrak{S}[\boldsymbol{\beta}_1, \cdots, \boldsymbol{\beta}_k]$究竟有一些什么性质。我们知道：如果$\mathfrak{R}$中的某子空间$\mathfrak{T}$含有向量$\boldsymbol{\beta}_1, \cdots, \boldsymbol{\beta}_k$，则由定理1易知$\boldsymbol{\beta}_1, \cdots, \boldsymbol{\beta}_k$的任意线性组合$c_1\boldsymbol{\beta}_1+\cdots+c_k\boldsymbol{\beta}_k$也在$\mathfrak{T}$内，因之有

$$\mathfrak{S}=\mathfrak{S}[\boldsymbol{\beta}_1, \cdots, \boldsymbol{\beta}_k]\subseteq\mathfrak{T}.$$

这是说：$\mathfrak{R}$中凡包含$\boldsymbol{\beta}_1, \cdots, \boldsymbol{\beta}_k$的任意子空间，都必包含$\mathfrak{S}$。而$\mathfrak{S}$又显然包含$\boldsymbol{\beta}_1, \cdots, \boldsymbol{\beta}_k$，故$\mathfrak{S}$为$\mathfrak{R}$中包含$\boldsymbol{\beta}_1, \cdots, \boldsymbol{\beta}_k$的子空间中最小的一个子空间。

当 $\boldsymbol{\beta}_1, \cdots, \boldsymbol{\beta}_k$ 线性无关时，$\boldsymbol{\beta}_1, \cdots, \boldsymbol{\beta}_k$ 就是 $\mathfrak{S}$ 的基底，因而这时 $\mathfrak{S}$ 的维数是 k；若 $\boldsymbol{\beta}_1, \cdots, \boldsymbol{\beta}_k$ 线性相关，则从其中除去那些能为其它的 $\boldsymbol{\beta}$ 所线性表示的那些 $\boldsymbol{\beta}$，而不失普遍性，可以令 $\boldsymbol{\beta}_1, \cdots, \boldsymbol{\beta}_l$ 为剩下的一组使再不能删掉了，因而 $\boldsymbol{\beta}_1, \cdots, \boldsymbol{\beta}_l$ 线性无关，且每 $\boldsymbol{\beta}_i (i=l+1, \cdots, k)$ 为 $\boldsymbol{\beta}_1, \cdots, \boldsymbol{\beta}_l$ 之线性组合[†]．于是，易知 $\mathfrak{S}=\mathfrak{S}[\boldsymbol{\beta}_1, \cdots, \boldsymbol{\beta}_k]=\mathfrak{S}[\boldsymbol{\beta}_1, \cdots, \boldsymbol{\beta}_l]$ 之基底为 $\boldsymbol{\beta}_1, \cdots, \boldsymbol{\beta}_l$，即 $\mathfrak{S}$ 是 l 维的空间．

由上面的讨论，可知向量空间 $\mathfrak{R}$ 的维数 m 不小于其子空间之维数，且子空间之维数等于 m 的充要条件是这子空间就是空间 $\mathfrak{R}$ 自身．

又，在 m 维向量空间 $\mathfrak{R}$ 的子空间中，有维数小于 m 的任意数的子空间：这是因为，若 $\boldsymbol{\alpha}_1, \cdots, \boldsymbol{\alpha}_m$ 为 $\mathfrak{R}$ 之基底，则 $\boldsymbol{\alpha}_1, \cdots, \boldsymbol{\alpha}_k (k \leqslant m)$ 当然线性无关，因之由 $\boldsymbol{\alpha}_1, \cdots, \boldsymbol{\alpha}_k$ 所生成的 $\mathfrak{R}$ 之子空间的维数为 k．当 $k<m$ 时，我们知道 $\boldsymbol{\alpha}_1, \cdots, \boldsymbol{\alpha}_m$ 中任 k 个都是线性无关的，因之在 m 维空间 $\mathfrak{R}$ 中维数为小于 m 之任意数的子空间不仅存在，而且还不止一个．而维数是 0 及 m 的子空间都只有一个，那就是零空间及 $\mathfrak{R}$ 自身．由第 2 节的定理 3，又可知 $\mathfrak{R}$ 的基底可以由它的子空间的基底去补充而得到．

下面，介绍 $\mathfrak{R}$ 的子空间的两个重要运算．

假定 $\mathfrak{R}_1', \cdots, \mathfrak{R}_t', \cdots$ 是向量空间 $\mathfrak{R}$ 的子空间，那末同时在所有子空间 $\mathfrak{R}_1', \cdots, \mathfrak{R}_t', \cdots$ 中一切向量的集合（也就是说，凡属于 $\mathfrak{R}_1', \cdots, \mathfrak{R}_t', \cdots$ 的公共向量的集合）是 $\mathfrak{R}$ 的子空间，叫做 $\mathfrak{R}_1', \cdots, \mathfrak{R}_t', \cdots$ 的交，用记号 $\mathfrak{R}_1' \cap \cdots \cap \mathfrak{R}_t' \cap \cdots$ 表示．这是

† 我们叫 $\boldsymbol{\beta}_1, \cdots, \boldsymbol{\beta}_l$ 为一组向量 $\boldsymbol{\beta}_1, \cdots, \boldsymbol{\beta}_k$ 中的**最大无关组**．易知任一组向量（有限多个或无限多个）之最大无关组不止一组，但任两个最大无关组所含向量的个数是相等的．

因为，假如 $\alpha, \beta \in \mathfrak{R}_1' \cap \cdots \cap \mathfrak{R}_t' \cap \cdots$，那末 $\alpha, \beta \in$ 每个 $\mathfrak{R}_i'$；因为 $\mathfrak{R}_i'$ 为子空间，故 $\alpha+\beta \in \mathfrak{R}_i'$，$\lambda\alpha \in \mathfrak{R}_i'$（$\lambda$ 是任意数），因此 $\alpha+\beta \in \mathfrak{R}_1' \cap \cdots \cap \mathfrak{R}_t' \cap \cdots$，$\lambda\alpha \in \mathfrak{R}_1' \cap \cdots \cap \mathfrak{R}_t' \cap \cdots$，所以 $\mathfrak{R}_1' \cap \cdots \cap \mathfrak{R}_t' \cap \cdots$ 是 $\mathfrak{R}$ 的子空间.

譬如，三维空间 $\mathfrak{R}_3$ 中，两个过坐标原点的不同平面的交，就是过原点的它们的交线.

除了子空间有交的运算外，还有和的运算. 具体地说，凡形状象 $\sum \alpha_i$（但和 $\sum$ 中只有有限多个项 α_i 不为零向量）的向量的集合是 $\mathfrak{R}$ 的子空间（$\alpha_i \in \mathfrak{R}_i'$），叫做 $\mathfrak{R}_1', \cdots, \mathfrak{R}_t', \cdots$ 的和，用记号 $\mathfrak{R}_1' \cup \mathfrak{R}_2' \cup \cdots \cup \mathfrak{R}_t' \cup \cdots$ 表示（这是因为，假如

$$\alpha = \sum \alpha_i, \quad \beta = \sum \beta_i \quad (\alpha_i, \beta_i \in \mathfrak{R}_i'),$$

但两个和 $\sum$ 中都只有有限多个项不为零（若存在这样的项），故 $\alpha+\beta$ 与 $\lambda\alpha$ 都可写成 $\sum \gamma_i (\gamma_i \in \mathfrak{R}_i')$ 的形状，且 $\sum$ 中只有有限多个 $\gamma_i \neq \mathbf{0}$，即 $\alpha+\beta$ 与 $\lambda\alpha$ 仍在 $\mathfrak{R}_1' \cup \mathfrak{R}_2' \cup \cdots \cup \mathfrak{R}_t' \cup \cdots$ 内，也就是说 $\mathfrak{R}_1' \cup \mathfrak{R}_2' \cup \cdots \cup \mathfrak{R}_t' \cup \cdots$ 为 $\mathfrak{R}$ 之子空间）.

容易验证，交 $\mathfrak{R}_1' \cap \mathfrak{R}_2' \cap \cdots \cap \mathfrak{R}_t' \cap \cdots$ 是包含在 $\mathfrak{R}_1', \mathfrak{R}_2', \cdots, \mathfrak{R}_t', \cdots$ 中的 $\mathfrak{R}$ 之最大子空间（这是因为，若 $\mathfrak{D}$ 为 $\mathfrak{R}$ 之子空间且 $\mathfrak{D} \subseteq \mathfrak{R}_i'$ $(i=1, \cdots, t, \cdots)$，则显然有 $\mathfrak{D} \subseteq \mathfrak{R}_1' \cap \mathfrak{R}_2' \cap \cdots \cap \mathfrak{R}_t' \cap \cdots$）.

类似地，易知和 $\mathfrak{R}_1' \cup \mathfrak{R}_2' \cup \cdots \cup \mathfrak{R}_t' \cup \cdots$ 是包含 $\mathfrak{R}_1', \mathfrak{R}_2', \cdots, \mathfrak{R}_t', \cdots$ 的 $\mathfrak{R}$ 之最小子空间.

关于两个子空间的维数，同它们的和与交之维数间，有下面一个重要关系：

定理 8 假定向量空间 $\mathfrak{R}$ 的子空间 $\mathfrak{R}_1', \mathfrak{R}_2'$ 的维数分别是 m_1, m_2，它们的和 $\mathfrak{R}_1' \cup \mathfrak{R}_2'$ 及交 $\mathfrak{R}_1' \cap \mathfrak{R}_2'$ 的维数分别是 n_1, n_2；那末 $n_1+n_2=m_1+m_2$.

证明 设 $\alpha_1, \cdots, \alpha_{n_2}$ 为 $\mathfrak{R}_1' \cap \mathfrak{R}_2'$ 之基底. 因 $\mathfrak{R}_1' \cap \mathfrak{R}_2'$ 是

$\mathfrak{R}_i'$ $(i=1, 2)$ 的子空间，故据第 2 节定理 3，可适当地挑选 $\boldsymbol{\beta}_1, \cdots, \boldsymbol{\beta}_k$ 及 $\boldsymbol{\gamma}_1, \cdots, \boldsymbol{\gamma}_l$，使

$$\boldsymbol{\alpha}_1, \cdots, \boldsymbol{\alpha}_{n_2}, \boldsymbol{\beta}_1, \cdots, \boldsymbol{\beta}_k$$

为 $\mathfrak{R}_1'$ 之基底；使

$$\boldsymbol{\alpha}_1, \cdots, \boldsymbol{\alpha}_{n_2}, \boldsymbol{\gamma}_1, \cdots, \boldsymbol{\gamma}_l$$

为 $\mathfrak{R}_2'$ 之基底．于是 $m_1=n_2+k$, $m_2=n_2+l$.

又，$\mathfrak{R}_1'\cup\mathfrak{R}_2'$ 之任意向量可写为 $\boldsymbol{\sigma}+\boldsymbol{\tau}$, $\boldsymbol{\sigma}\in\mathfrak{R}_1'$, $\boldsymbol{\tau}\in\mathfrak{R}_2'$, 因 $\boldsymbol{\sigma}$ 为 $\boldsymbol{\alpha}_1, \cdots, \boldsymbol{\alpha}_{n_2}, \boldsymbol{\beta}_1, \cdots, \boldsymbol{\beta}_k$ 之线性组合，$\boldsymbol{\tau}$ 为 $\boldsymbol{\alpha}_1, \cdots, \boldsymbol{\alpha}_{n_2}, \boldsymbol{\gamma}_1, \cdots, \boldsymbol{\gamma}_l$ 之线性组合，故 $\boldsymbol{\sigma}+\boldsymbol{\tau}$ 为 $\boldsymbol{\alpha}_1, \cdots, \boldsymbol{\alpha}_{n_2}, \boldsymbol{\beta}_1, \cdots, \boldsymbol{\beta}_k, \boldsymbol{\gamma}_1, \cdots, \boldsymbol{\gamma}_l$ 之线性组合．另方面，若

$$a_1\boldsymbol{\alpha}_1+\cdots+a_{n_2}\boldsymbol{\alpha}_{n_2}+b_1\boldsymbol{\beta}_1+\cdots+b_k\boldsymbol{\beta}_k+c_1\boldsymbol{\gamma}_1+\cdots+c_l\boldsymbol{\gamma}_l=\mathbf{0}, \quad (3)$$

则 $$c_1\boldsymbol{\gamma}_1+\cdots+c_l\boldsymbol{\gamma}_l=-\sum_{i=1}^{n_2}a_i\boldsymbol{\alpha}_i-\sum_{i=1}^{k}b_i\boldsymbol{\beta}_i\in\mathfrak{R}_1'\cap\mathfrak{R}_2',$$

故 $$c_1\boldsymbol{\gamma}_1+\cdots+c_l\boldsymbol{\gamma}_l=a_1'\boldsymbol{\alpha}_1+\cdots+a_{n_2}'\boldsymbol{\alpha}_{n_2},$$

再把它代入(3)中，得

$$(a_1+a_1')\boldsymbol{\alpha}_1+\cdots+(a_{n_2}+a_{n_2}')\boldsymbol{\alpha}_{n_2}+b_1\boldsymbol{\beta}_1+\cdots+b_k\boldsymbol{\beta}_k=\mathbf{0}.$$

于是，从 $\mathfrak{R}_1'$ 之基底说来，就有 $b_1=\cdots=b_k=0$，因而(3)式又变为 $a_1\boldsymbol{\alpha}_1+\cdots+a_{n_2}\boldsymbol{\alpha}_{n_2}+c_1\boldsymbol{\gamma}_1+\cdots+c_l\boldsymbol{\gamma}_l=\mathbf{0}$；再从 $\mathfrak{R}_2'$ 之基底说来，又必有 $a_1=\cdots=a_{n_2}=c_1=\cdots=c_l=0$．这已说明了 $\boldsymbol{\alpha}_1, \cdots, \boldsymbol{\alpha}_{n_2}; \boldsymbol{\beta}_1, \cdots, \boldsymbol{\beta}_k, \boldsymbol{\gamma}_1, \cdots, \boldsymbol{\gamma}_l$ 为 $\mathfrak{R}_1'\cup\mathfrak{R}_2'$ 之一组基底，故必 $n_1=n_2+k+l$, $n_1+n_2=n_2+k+l+n_2=m_1+m_2$．定理 3 获证．

练 习 十 二

1．证明下列各种向量的集合是 $\mathfrak{R}_n$ 的子空间，并说明在直觉空间($n=2, 3$)中是怎样的图象．

(1) $\mathfrak{R}_n$ 之向量，其第一个分量为 0；

(2) $\mathfrak{R}_n$ 之向量 $(x_1, x_2, \cdots, x_n)$，其第一、二两个分量的和等于零(即 $x_1+x_2=0$)．

2. 设 $\mathfrak{S}_1$ 及 $\mathfrak{S}_2$ 为向量空间 $\mathfrak{R}$ 的两个子空间。证明：$\mathfrak{S}_1\cup\mathfrak{S}_2$ 中向量 $\boldsymbol{\alpha}$ 写成 $\boldsymbol{\alpha}=\boldsymbol{\alpha}_1+\boldsymbol{\alpha}_2, \boldsymbol{\alpha}_1\in\mathfrak{S}_1, \boldsymbol{\alpha}_2\in\mathfrak{S}_2$ 的表示式是唯一的充要条件是 $\mathfrak{S}_1\cap\mathfrak{S}_2=\mathfrak{O}$，这里 $\mathfrak{O}$ 表示零空间.

3. 设 $\mathfrak{S}_1$ 与 $\mathfrak{S}_2$ 是 $\mathfrak{R}_n$ 的两个子空间，它们的维数都 $<n$，证明 $\mathfrak{R}_n$ 中必存在一个既不属于 $\mathfrak{S}_1$ 又不属于 $\mathfrak{S}_2$ 的向量.

4. 证明：设 n 维空间 $\mathfrak{R}_n$ 的两个线性子空间之和的维数比其交的维数多 1，则和必与其中一个子空间重合，交又与另一个子空间重合.

第 4 节　线性方程组的解的几何意义

在直觉的三维向量空间中，有直线、平面这些基本概念. 下面，我们来推广这概念.

我们知道：假如 $\boldsymbol{x}_0$ 是三维向量空间 $\mathfrak{R}_3$ 的向量，那末所有形如

$$\boldsymbol{x}=\boldsymbol{x}_0+\lambda\boldsymbol{y}\quad(\lambda\text{ 是任意数})$$

的向量的全体，构成 $\mathfrak{R}_3$ 中过点 $\boldsymbol{x}_0$ 的直线，这里 $\boldsymbol{y}$ 为 $\mathfrak{R}_3$ 中一个固定的向量. 又，所有形如

$$\boldsymbol{x}=\boldsymbol{x}_0+\lambda\boldsymbol{y}+\mu\boldsymbol{z}\quad(\lambda,\ \mu\text{ 为任意数})$$

的向量的全体，构成 $\mathfrak{R}_3$ 中过点 $\boldsymbol{x}_0$ 的平面，这里 $\boldsymbol{y}, \boldsymbol{z}$ 是 $\mathfrak{R}_3$ 中两个固定的线性无关的向量.

现在，再推广到一般 n 维空间 $\mathfrak{R}_n$ 中.

假如 $\boldsymbol{x}_0$ 是 $\mathfrak{R}_n$ 中的向量，那末所有形状象

$$\boldsymbol{x}=\boldsymbol{x}_0+\boldsymbol{y}$$

的向量的全体，当 $\boldsymbol{y}$ 为 $\mathfrak{R}_n$ 的一维子空间中任意向量时，叫数 $\mathfrak{R}_n$ 的直线，而当 $\boldsymbol{y}$ 为 $\mathfrak{R}_n$ 的二维子空间中任意向量时，叫做 $\mathfrak{R}_n$ 的平面.

由第二章，线性非齐次方程组的解向量 $\boldsymbol{x}$，可以写成它的一个固定的解向量 $\boldsymbol{x}_0$ 与它对应的齐次方程组的解向量 $\boldsymbol{y}$ 的

和，又 n 个未知量秩是 r 的线性齐次方程组的解向量空间的维数是 $n-r$，于是，秩为 $n-1$ 的线性方程组

$$\begin{cases} a_{11}x_1+\cdots+a_{1n}x_n=b_1, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{n-1,1}x_1+\cdots+a_{n-1,n}x_n=b_{n-1} \end{cases}$$

的解向量的全体，形成 $\mathfrak{R}_n$ 的一直线，秩是 $n-2$ 的线性方程组

$$\begin{cases} a_{11}x_1+\cdots+a_{1n}x_n=b_1, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{n-2,1}x_1+\cdots+a_{n-2,n}x_n=b_{n-2} \end{cases}$$

的解向量的全体，形成 $\mathfrak{R}_n$ 的一个平面．这与 $\mathfrak{R}_3$ 中平面用一个线性方程表示，直线用两个线性方程表示的情况相一致．

至于线性方程组(齐次或非齐次的)之基础解系的有关结果，其几何意义又怎样呢？有些已经在上面谈了，譬如 n 个未知量的线性齐次方程组的秩是 r 的时候，它的一切解向量的集合是 $\mathfrak{R}_n$ 的一个 $n-r$ 维的子空间．

现在来讨论非齐次的线性方程组．已知

$$\begin{cases} a_{11}x_1+\cdots+a_{1n}x_n=b_1, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{m1}x_1+\cdots+a_{mn}x_n=b_m. \end{cases} \tag{4}$$

在系数矩阵与增广矩阵有等秩 r 时是有解的．作(4)之相应的齐次方程组

$$\begin{cases} a_{11}x_1+\cdots+a_{1n}x_n=0, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{m1}x_1+\cdots+a_{mn}x_n=0, \end{cases} \tag{5}$$

并取(4)之某一个固定解 $P(a_1, a_2, \cdots, a_n)$ 后，则(4)之任意解 $Q(x_1, x_2, \cdots, x_n)$ 必定是而且也只能是

$$(x_1, x_2, \cdots, x_n)=(a_1, a_2, \cdots, a_n)+(c_1, c_2, \cdots, c_n)$$

的形状，其中 $\boldsymbol{r}=(c_1, c_2, \cdots, c_n)$ 为(5)的任意解向量．这是

说：向量 $\overrightarrow{PQ}=\{x_1-a_1,\ x_2-a_2,\ \cdots,\ x_n-a_n\}=\boldsymbol{r}$，它的几何意义是：(4)之一切解点 $Q(x_1,\ \cdots,\ x_n)$（当然也包括 P 点在内）都可以用下述方法来求得：从 P 点引等于(5)之 $n-r$ 维解空间 $\mathfrak{L}$ 中的一切向量所得到的终点．这就是将空间 $\mathfrak{L}$ 施行平行移动后即可得到(4)的一切解点．这与直觉空间中 ($n=2,\ 3$ 时)的情况是一致的：譬如说，二维空间 $\mathfrak{R}_2$ 中的直线 l：$ax+by+c=0(c\neq0)$ 恰与直线 l_0：$ax+by=0$ 平行；又如三维空间 $\mathfrak{R}_3$ 中的平面 $ax+by+cz+d=0$（其中 $d\neq0$），也恰与平面 $ax+by+cz=0$平行；$\mathfrak{R}_3$ 中的直线

$$\begin{cases}a_1x+b_1y+c_1z+d_1=0,\\ a_2x+b_2y+c_2z+d_2=0\end{cases}$$

与直线
$$\begin{cases}a_1x+b_1y+c_1z=0,\\ a_2x+b_2y+c_2z=0\end{cases}$$

也是平行的．

最后，关于基础解系的几何意义，想多说几句话．从直觉空间($n=2,\ 3$)开始．我们知道，二维空间 $\mathfrak{R}_2$ 中的直线 l 得由它上面两个不同的点 $M_1(x_1,\ y_1)$与 $M_2(x_2,\ y_2)$来决定，因而 l 的方程可象推导出(1)式的那样(如图 3–1)，得参数式

$$\begin{cases}x=x_1+\lambda(x_2-x_1),\\ y=y_1+\lambda(y_2-y_1).\end{cases}$$

将上式中 $x=x_1+\lambda(x_2-x_1)=(1-\lambda)x_1+\lambda x_2$ 的 $1-\lambda$ 写为 μ，即 $x=\mu x_1+\lambda x_2$，则 $\mu+\lambda=1$．同样，有 $y=\mu y_1+\lambda y_2$．这说明了直线 l 得由它上面两个不同的点 $M_1(x_1,\ y_1)$与 $M_2(x_2,\ y_2)$ 用参数方程

$$\begin{cases}x=\lambda x_1+\mu x_2,\\ y=\lambda y_1+\mu y_2\end{cases}\tag{6}$$

来表示，但 $\lambda+\mu=1$. 另外，l 又可由 x, y 之一次方程 $ax+by+c=0$ 来表示. 并且，当且仅当 l 不通过坐标原点时，(x_1, y_1) 与 (x_2, y_2) 才是线性无关的，这时 $ax+by+c=0$ 为非齐次的 $(c\neq0)$. 当且仅当 l 通过原点时，$ax+by+c=0$ 是齐次的，即 $c=0$，因而这时 (x_1, y_1) 与 (x_2, y_2) 成比例（线性相关），随之(6)式可简化为

$$\begin{cases} x=\nu x_1, \\ y=\nu y_1. \end{cases} \tag{7}$$

综上所述，就得到：**对于 x, y 的一个齐次线性方程，只有一个线性无关的解 (x_1, y_1)，使得任何解 (x, y) 都和它成比例. 而对于 x, y 的一个非齐次线性方程，有两个线性无关的解 (x_1, y_1) 与 (x_2, y_2)，使得任何解 (x, y) 能表写为它们的线性组合，但组合系数之和等于 1.**

用几何语言说明，就是：**二维空间 $\mathfrak{R}_2$ 中，通过坐标原点的直线上任一点的坐标，可以表成该直线上异于原点的任一所给点的坐标的倍数. 不通过坐标原点的直线上任一点的坐标，可写成直线上任两个所给点的坐标的线性组合（这两个点必为线性无关的），且组合之系数的和等于 1.**

注意：所谓直线是一维向量空间，指的是它只有一个线性无关的向量，这里，直线是要通过坐标原点的，这是因为向量空间必定含零向量（几何中的向量都是以原点为起点的）. 而不通过坐标原点的直线 l 虽习惯上也叫做一维空间，但它已经不是向量空间了†. 所以它上面两个线性无关的点 M_1, M_2 所对应的两个线性无关的向量 $\overrightarrow{OM}_1$ 和 $\overrightarrow{OM}_2$ 已经不在这直线 l 上，如图 3-6 所示. 因之不要误解直线 l 有两个无关的点就有两个无关的向量，实际上，l 上的向量中只有一个无

† 向量空间也叫做线性空间.

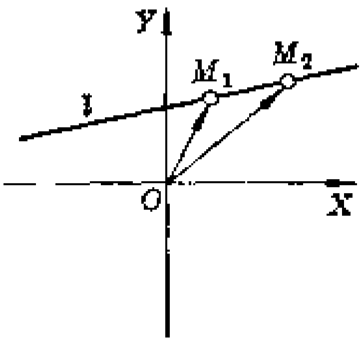


图 3-6

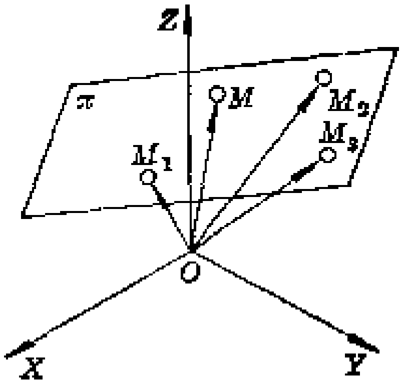


图 3-7

关的向量，这时向量的起点是放在 l 上的.

对三维空间 $\mathfrak{R}_3$ 来讲，平面 π 由它上面不共线的三个点 M_1, M_2, M_3 来决定(图 3-7)，令 M_i 点之坐标为 (x_i, y_i, z_i)，$(i=1, 2, 3)$. 在平面 π 上取任意一点 $M(x, y, z)$，则因

$$\overrightarrow{M_1M}=\mu\overrightarrow{M_1M_2}+\nu\overrightarrow{M_1M_3},$$

故

$$\begin{aligned}\{x, y, z\}&=\overrightarrow{OM}=\overrightarrow{OM_1}+\overrightarrow{M_1M}=\overrightarrow{OM_1}+\mu\overrightarrow{M_1M_2}+\nu\overrightarrow{M_1M_3}\\&=\{x_1, y_1, z_1\}+\mu\{x_2-x_1, y_2-y_1, z_2-z_1\}\\&\quad+\nu\{x_3-x_1, y_3-y_1, z_3-z_1\}\\&=\{x_1+\mu(x_2-x_1)+\nu(x_3-x_1), y_1+\mu(y_2-y_1)\\&\quad+\nu(y_3-y_1), z_1+\mu(z_2-z_1)+\nu(z_3-z_1)\}\\&=\{\lambda x_1+\mu x_2+\nu x_3, \lambda y_1+\mu y_2+\nu y_3, \lambda z_1+\mu z_2+\nu z_3\},\end{aligned}$$

式中 $\lambda=1-\mu-\nu$. 于是，

$$\begin{cases}x=\lambda x_1+\mu x_2+\nu x_3,\\y=\lambda y_1+\mu y_2+\nu y_3,\\z=\lambda z_1+\mu z_2+\nu z_3.\end{cases}\qquad(8)$$

这说明了平面 π 可由它上面不共线的三点 $M_i(x_i, y_i, z_i)$，$(i=1, 2, 3)$，用参数方程(8)来表示，其中 $\lambda+\mu+\nu=1$. 另一方面，平面 π 又可由 x, y, z 的一次方程来表示，如

$$ax+by+cz+d=0.$$

并且，当且仅当平面 π 不通过坐标原点时，(x_1, y_1, z_1)，(x_2, y_2, z_2) 与 (x_3, y_3, z_3) 才是线性无关的，这时 $d\neq0$ 即 $ax+by+cz+d=0$ 为非齐次的. 当且仅当 π 通过原点时，(x_1, y_1, z_1)，(x_2, y_2, z_2) 与 (x_3, y_3, z_3) 才线性相关，而 $ax+by+cz+d=0$ 是齐次的，即 $d=0$，这时(8)式可简化为

$$\begin{cases} x=\sigma x_1+\tau x_2, \\ y=\sigma y_1+\tau y_2, \\ z=\sigma z_1+\tau z_2. \end{cases} \tag{9}$$

综上所述，就得到：*三维空间 $\mathfrak{R}_3$ 中通过坐标原点的平面上的任一点的坐标，可以表成它上面与原点不同在一直线上的任两个所给点的坐标的线性组合. 不通过坐标原点的平面上任一点的坐标，可表成平面上任三个不共线的所给点(因之一定线性无关)的坐标的线性组合，且组合之系数的和等于 1.*

注意：不通过坐标原点 O 的平面 π (如图 3-7)，习惯上虽叫做二维空间，但它已不是线性空间，因它不包含零向量(即坐标原点)；在它上面三个无关的点 M_1, M_2, M_3 虽对应着三个线性无关的向量 $\overrightarrow{OM}_1$, $\overrightarrow{OM}_2$, $\overrightarrow{OM}_3$，但这三个向量并不在这平面上，而这平而上只有两个线性无关的向量，这时向量的起点是放在 π 上的.

推广这个概念，得到线性方程组(齐次的与非齐次的)之几何解释于下：*秩为 r 之齐次线性方程组的一切解点的集合形成一个 $n-r$ 维的向量空间，是 $\mathfrak{R}_n$ 的子空间. 秩为 r 之非齐次线性方程组的一切解点的集合是 $\mathfrak{R}_n$ 内的一个 $n-r$ 空间(但不是线性空间或向量空间)，其中有 $n-r+1$ 个线性无关的点(它们对应着的 $n-r+1$ 个无关的向量并不在这空间内)，使得这空间内任一点可表成它们的线性组合，且组合系数的和等于 1.*

练 习 十 三

1. 对于 $\mathfrak{R}_n$ 之任一个 k 维子空间 $\mathfrak{L}$，必可作一组含 n 个未知量的线性齐次方程，使 $\mathfrak{L}$ 就是这方程组的解向量空间.

2. 设 $\boldsymbol{A}, \boldsymbol{B}$ 都是 n 级矩阵，则 $\boldsymbol{AB}$ 的秩 $=\boldsymbol{B}$ 的秩的充要条件是方程组 $\boldsymbol{ABX}=\boldsymbol{0}$ 的解必为方程组 $\boldsymbol{BX}=\boldsymbol{0}$ 的解.

3. 设 $\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C}$ 都是 n 级矩阵，若 $r_{\boldsymbol{A}}=r_{\boldsymbol{BA}}$，则 $r_{\boldsymbol{AC}}=r_{\boldsymbol{BAC}}$.

4. 设 $\boldsymbol{A}$ 为一方阵，k 是一正整数使 $r_{\boldsymbol{A}^k}=r_{\boldsymbol{A}^{k+1}}$，试证 $r_{\boldsymbol{A}^{k+1}}=r_{\boldsymbol{A}^{k+2}}=\cdots$.

第5节 向量空间定义的公理化

在第1节已介绍了 n 维向量空间或其子空间的概念，它是 n 元向量之集合，它具有两种运算：即其中任两个向量相加也仍在这集合内；其中任一向量与任意数之积也在这集合内. 这两种运算还具备一些基本性质. 我们知道，数学中还讨论一些别的对象，它们不一定都能象用由 n 个数组所构成的 n 元向量来表示，但却有与上面类似的两种运算，并且在它们所成之集合中也具有与上面类似关于运算的基本性质. 为了普遍地研究，有必要把这样对象不同的集合从本质上统一起来，这就可以舍弃具体的对象，而根据运算的基本性质，用公理的形式引进向量空间的一般定义.

定义 1 设有一集合 $\mathfrak{V}$，它的元叫做"向量"，用字母 $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$, $\boldsymbol{\gamma}$, $\cdots$ $\boldsymbol{x}$, $\boldsymbol{y}$, $\boldsymbol{z}$, $\cdots$ 来表示. 如果 $\mathfrak{V}$ 满足下列各条件，就叫 $\mathfrak{V}$ 为 n 维向量空间：

i) $\mathfrak{V}$ 中有叫"加法"及"数积"的这样两种运算，并且还有"加法"的逆运算"减法"，即如果 $\boldsymbol{\alpha}, \boldsymbol{\beta}\in\mathfrak{V}$，那末 $\boldsymbol{\alpha}+\boldsymbol{\beta}\in\mathfrak{V}$，$k\boldsymbol{\alpha}\in\mathfrak{V}$（$k$ 为任意数），并且还有元 $\boldsymbol{x}\in\mathfrak{V}$ 使 $\boldsymbol{\alpha}+\boldsymbol{x}=\boldsymbol{\beta}$；

ii) $\mathfrak{V}$ 中"加法"满足交换律、结合律；

$$\alpha+\beta=\beta+\alpha,$$
$$(\alpha+\beta)+\gamma=\alpha+(\beta+\gamma),$$

"数积"满足分配律、结合律:

$$k(\alpha+\beta)=k\alpha+k\beta,$$
$$(k+l)\alpha=k\alpha+l\alpha,$$
$$(kl)\alpha=k(l\alpha);$$

iii) $\mathfrak{V}$ 中有 n 个固定的向量 $u_1, u_2, \cdots, u_n$, 使 $\mathfrak{V}$ 中每个向量 α 都可用唯一的方法表成它们的线性组合:

$$\alpha=k_1u_1+k_2u_2+\cdots+k_nu_n.$$

n 维向量空间又称为 n 维线性空间, 或 n 维仿射空间.

$u_1, u_2, \cdots, u_n$ 称为 $\mathfrak{V}$ 之基底, 而 $k_1, k_2, \cdots, k_n$ 这 n 个数称为向量 α 关于基底 $u_1, u_2, \cdots, u_n$ 的坐标.

定义 1 中的 k, l 都是数, 如果这些数都限制在实数的范围内, 就称 $\mathfrak{V}$ 为实空间; 如果允许取复数, 就叫 $\mathfrak{V}$ 为复空间.

注意: 线性空间是一个抽象名词, 它的元一般也是抽象的, 当然不一定是数; 所谓实空间只是其中的系数 k、l 等都限制取实数而已, 其实在实空间中一般并不包含一个实数, 就是复空间同样也不包含一个复数.

下面再叙述线性空间的一些基本性质.

性质 1 对每个 $\alpha\in\mathfrak{V}$, 恒有 $1\cdot\alpha=\alpha$.

证明
$$\begin{aligned}1\cdot\alpha&=1\cdot(k_1u_1+k_2u_2+\cdots+k_nu_n)\\&=1\cdot(k_1u_1)+1\cdot(k_2u_2)+\cdots+1\cdot(k_nu_n)\\&=(1\cdot k_1)u_1+(1\cdot k_2)u_2+\cdots+(1\cdot k_n)u_n\\&=k_1u_1+k_2u_2+\cdots+k_nu_n=\alpha.\end{aligned}$$

性质 2 $\mathfrak{V}$ 中有"零"元 $\mathbf{0}$, 即对任 $\alpha\in\mathfrak{V}$ 恒有 $\alpha+\mathbf{0}=\mathbf{0}+\alpha=\alpha$; 且零元 $\mathbf{0}$ 还是唯一的.

证明 设 $\alpha\in\mathfrak{V}$, 由定义 1 中的条件(i), 知有元 $x\in\mathfrak{V}$ 满

足 $\boldsymbol{\alpha}+\boldsymbol{x}=\boldsymbol{\alpha}$. 于是，对任意的 $\boldsymbol{\beta}\in\mathfrak{V}$，由于有 $\boldsymbol{y}\in\mathfrak{V}$ 使 $\boldsymbol{\alpha}+\boldsymbol{y}=\boldsymbol{\beta}$，故有

$$\boldsymbol{\beta}+\boldsymbol{x}=(\boldsymbol{\alpha}+\boldsymbol{y})+\boldsymbol{x}=(\boldsymbol{\alpha}+\boldsymbol{x})+\boldsymbol{y}=\boldsymbol{\alpha}+\boldsymbol{y}=\boldsymbol{\beta}.$$

这就说明了 $\mathfrak{V}$ 中任意元 $\boldsymbol{\beta}$ 与元 $\boldsymbol{x}$ 的和仍为 $\boldsymbol{\beta}$，即 $\boldsymbol{x}$ 得为性质 2 中所说的"零"元的作用，就以 $\mathbf{0}$ 记之，即 $\boldsymbol{x}=\mathbf{0}$. 若尚有 $\mathbf{0}'\in\mathfrak{V}$ 使对任 $\boldsymbol{\alpha}\in\mathfrak{V}$ 恒有 $\boldsymbol{\alpha}+\mathbf{0}'=\boldsymbol{\alpha}$，则取 $\boldsymbol{\alpha}=\mathbf{0}$ 时就得到 $\mathbf{0}+\mathbf{0}'=\mathbf{0}$; 但又由于对任 $\boldsymbol{\alpha}\in\mathfrak{V}$ 恒有 $\boldsymbol{\alpha}+\mathbf{0}=\boldsymbol{\alpha}$, 故特取 $\boldsymbol{\alpha}=\mathbf{0}'$ 就得知 $\mathbf{0}'+\mathbf{0}=\mathbf{0}'$；于是由于交换律而应有 $\mathbf{0}+\mathbf{0}'=\mathbf{0}'+\mathbf{0}$，故得 $\mathbf{0}'=\mathbf{0}$. 这就证明了"零"元的唯一性, 证完.

性质 3 $\mathfrak{V}$ 中每元有唯一个"负"元 (即对每 $\boldsymbol{\alpha}\in\mathfrak{V}$，有 $-\boldsymbol{\alpha}\in\mathfrak{V}$ 使 $\boldsymbol{\alpha}+(-\boldsymbol{\alpha})=\mathbf{0}$).

证明 据定义 1 中条件(i)，知有元 $\boldsymbol{x}\in\mathfrak{V}$ 使 $\boldsymbol{\alpha}+\boldsymbol{x}=\mathbf{0}$，这元 $\boldsymbol{x}$ 就担负了元 $\boldsymbol{\alpha}$ 之"负"元 $-\boldsymbol{\alpha}$ 的作用. 如果除上述 $\boldsymbol{x}$ 外尚有元 $\boldsymbol{x}'\in\mathfrak{V}$ 使 $\boldsymbol{\alpha}+\boldsymbol{x}'=\mathbf{0}$(当然, $\boldsymbol{x}$ 及 $\boldsymbol{x}'$ 是随 $\boldsymbol{\alpha}$ 而变的)，则

$$\boldsymbol{x}'=\mathbf{0}+\boldsymbol{x}'=(\boldsymbol{\alpha}+\boldsymbol{x})+\boldsymbol{x}'=(\boldsymbol{\alpha}+\boldsymbol{x}')+\boldsymbol{x}=\mathbf{0}+\boldsymbol{x}=\boldsymbol{x}.$$

这就证明了"负"元的唯一性, 证完.

性质 4 方程 $\boldsymbol{\alpha}+\boldsymbol{x}=\boldsymbol{\beta}$ 的解 $\boldsymbol{x}$ 是唯一的.

证明 若尚有 $\boldsymbol{\alpha}+\boldsymbol{x}'=\boldsymbol{\beta}(=\boldsymbol{\alpha}+\boldsymbol{x})$，则

$$\begin{aligned}\boldsymbol{x}'&=\mathbf{0}+\boldsymbol{x}'=((-\boldsymbol{\alpha})+\boldsymbol{\alpha})+\boldsymbol{x}'=(-\boldsymbol{\alpha})+(\boldsymbol{\alpha}+\boldsymbol{x}')\\&=(-\boldsymbol{\alpha})+(\boldsymbol{\alpha}+\boldsymbol{x})=((-\boldsymbol{\alpha})+\boldsymbol{\alpha})+\boldsymbol{x}=\mathbf{0}+\boldsymbol{x}=\boldsymbol{x}.\end{aligned}$$

证完.

性质 5 $0\cdot\boldsymbol{\alpha}=\mathbf{0}$(任 $\boldsymbol{\alpha}\in\mathfrak{V}$).

证明 $0\cdot\boldsymbol{\alpha}=(0+0)\cdot\boldsymbol{\alpha}=0\cdot\boldsymbol{\alpha}+0\cdot\boldsymbol{\alpha}$，故由性质 4，即得 $0\cdot\boldsymbol{\alpha}=\mathbf{0}$.

性质 6 对任何数 k，恒有 $k\cdot\mathbf{0}=\mathbf{0}$.

证明 $k\mathbf{0}=k(\mathbf{0}+\mathbf{0})=k\mathbf{0}+k\mathbf{0}$，故据性质 4，得 $k\mathbf{0}=\mathbf{0}$.

性质 7 $-\boldsymbol{a}=(-1)\boldsymbol{a}$.

证明 $\boldsymbol{a}+(-1)\boldsymbol{a}=1\cdot\boldsymbol{a}+(-1)\boldsymbol{a}=(1+(-1))\boldsymbol{a}=0\boldsymbol{a}=\boldsymbol{0}$, 故由"负"元之唯一性, 有 $(-1)\boldsymbol{a}=-\boldsymbol{a}$.

定义 1 实际上是一种公理化的方法，其目的在于考虑适合这些公理的集合的共同性质. 因为是研究它们运算的共同性质, 所以没有必要考虑所给集合里面对象的具体特征, 这就使问题的研究结果带有普遍化意义，而不致局限于各别的对象了. 当然, 上述公理化之定义较为抽象, 可能使初学者有不易捉摸之感, 于是自然会提出这样一个问题, 即上述这样纯公理化所定义的 n 维向量空间 $\mathfrak{V}$ 和前几节所讨论的具体的 n 维向量空间 $\mathfrak{R}_n$ 之间究竟有怎样的联系呢?

为了解决这个问题, 先引进一一对应和同构这两个概念.

定义 2 设有两个向量空间 $\mathfrak{S}$ 与 $\mathfrak{S}'$. 如果有一个明确的方法能够使 $\mathfrak{S}$ 的任一向量 $\boldsymbol{x}$ 和 $\mathfrak{S}'$ 的某一向量 $\boldsymbol{x}'$ 对应, 而且这对应关系又满足下面两个条件:

i) 对于每 $\boldsymbol{x}\in\mathfrak{S}$, 有一个且仅一个 $\boldsymbol{x}'\in\mathfrak{S}'$ 与它对应;

ii) 对每 $\boldsymbol{x}'\in\mathfrak{S}'$, 有一个且仅一个 $\boldsymbol{x}\in\mathfrak{S}$ 被它对应,

这时, 就叫 $\mathfrak{S}$ 与 $\mathfrak{S}'$ 的元间成一一对应的.

定义 3 如果在 $\mathfrak{S}$ 和 $\mathfrak{S}'$ 的向量之间能够建立一个这样一一对应的关系, 满足

i) 若向量 $\boldsymbol{x}\in\mathfrak{S}$ 与 $\boldsymbol{x}'\in\mathfrak{S}'$ 对应 (今后用符号 $\boldsymbol{x}\leftrightarrow\boldsymbol{x}'$ 表示), 那末不论 λ 为任何数, 恒有 $\lambda\boldsymbol{x}\leftrightarrow\lambda\boldsymbol{x}'$;

ii) 若 $\boldsymbol{x}\leftrightarrow\boldsymbol{x}'$, $\boldsymbol{y}\leftrightarrow\boldsymbol{y}'$, 那末 $\boldsymbol{x}+\boldsymbol{y}\leftrightarrow\boldsymbol{x}'+\boldsymbol{y}'$,

就叫空间 $\mathfrak{S}$ 和空间 $\mathfrak{S}'$ 同构, 记成 $\mathfrak{S}\cong\mathfrak{S}'$.

由定义 3 可知, 两个空间同构时, 这两个空间所不同者只是它们的元素间的记号不同，而由它们互相对应之元间由运算所产生的结果也必互相对应，这就说明了两个同构的空间

并没有本质上的差异(如果不问各个空间之元的具体特征,而只注意空间中"加法"与"数积"这两种运算).

根据同构的意义,可以证明重要的

定理 1 *任意两个 n 维向量空间 $\mathfrak{B}$ 和 $\mathfrak{B}'$ 一定是同构的.*

证明 假定 $\boldsymbol{u}_1, \boldsymbol{u}_2, \cdots, \boldsymbol{u}_n$ 和 $\boldsymbol{u}'_1, \boldsymbol{u}'_2, \cdots, \boldsymbol{u}'_n$ 分别是 $\mathfrak{B}$ 和 $\mathfrak{B}'$ 的基底. 我们在 $\mathfrak{B}$ 和 $\mathfrak{B}'$ 的元间建立下面的对应关系:

$$\boldsymbol{x} = \sum_{k=1}^{n} c_k \boldsymbol{u}_k \leftrightarrow \boldsymbol{x}' = \sum_{k=1}^{n} c_k \boldsymbol{u}'_k$$

(即当 $\mathfrak{B}$ 之向量 $\boldsymbol{x}$ 对于基底 $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_n$ 之坐标和 $\mathfrak{B}'$ 之向量 $\boldsymbol{x}'$ 对于基底 $\boldsymbol{u}'_1, \boldsymbol{u}'_2, \cdots, \boldsymbol{u}'_n$ 之坐标相同的时候,就令 $\boldsymbol{x} \leftrightarrow \boldsymbol{x}'$). 显然, 这种对应关系是一一对应的. 现在再来证明它满足同构条件:

事实上,假定在 $\mathfrak{B}$ 中任取两元

$$\boldsymbol{x} = \sum_{k=1}^{n} c_k \boldsymbol{u}_k \quad 与 \quad \boldsymbol{y} = \sum_{k=1}^{n} d_k \boldsymbol{u}_k,$$

则在 $\mathfrak{B}'$ 中与它们对应的元应分别是

$$\boldsymbol{x}' = \sum_{k=1}^{n} c_k \boldsymbol{u}'_k \quad 与 \quad \boldsymbol{y}' = \sum_{k=1}^{n} d_k \boldsymbol{u}'_k.$$

由于

$$\begin{aligned}\boldsymbol{x} + \boldsymbol{y} = \sum_{k=1}^{n} (c_k + d_k) \boldsymbol{u}_k \leftrightarrow (\boldsymbol{x} + \boldsymbol{y})' &= \sum_{k=1}^{n} (c_k + d_k) \boldsymbol{u}'_k \\ &= \sum_{k=1}^{n} c_k \boldsymbol{u}'_k + \sum_{k=1}^{n} d_k \boldsymbol{u}'_k \\ &= \boldsymbol{x}' + \boldsymbol{y}',\end{aligned}$$

以及

$$\begin{aligned}\lambda \boldsymbol{x} = \sum_{k=1}^{n} (\lambda c_k) \boldsymbol{u}_k \leftrightarrow (\lambda \boldsymbol{x})' &= \sum_{k=1}^{n} (\lambda c_k) \boldsymbol{u}'_k \\ &= \lambda \cdot \sum_{k=1}^{n} c_k \boldsymbol{u}'_k = \lambda \boldsymbol{x}',\end{aligned}$$

就又验证了同构的条件, 证完.

由定理 1，即得：

推论 任何 n 维向量空间一定与以往所说的 $\mathfrak{R}_n$ 同构.

正因为有这推论，所以抽象的(即公理化的) n 维向量空间和具体的 $\mathfrak{R}_n$ 就运算性质说来是没有什么本质上的差别的，我们之所以开始讲 $\mathfrak{R}_n$ 的道理也在这里，并还显得具体一些.

有了同构的概念，很多问题可以整体来讨论，不必各别研究. 凡属同构的，只要取其中一个来讨论，得到的结果对于其中任一个都能适合. 这样，很多问题只是转化为是否同构的问题了. 例如，由定义即知：一个变量 x 的所有小于 n 次多项式的集合是一个 n 维向量空间，1, x, …, …, x^{n-1} 是它的一组基底. 如果只研究多项式的加法以及它们与数的乘法间的性质，那末由定理 1 则知它和前几节说的 $\mathfrak{R}_n$ 同构，当明白了 $\mathfrak{R}_n$ 的构造后，也就明白了小于 n 次的一切多项式的集合之构造，不必再来推导了.

下面再证明定理 1 的逆定理，即：

定理 2 若向量空间 $\mathfrak{B}$ 与 $\mathfrak{B}'$ 是同构的，则它们的维数相等.

证明 设 $\mathfrak{B}$ 是 n 维的，$\boldsymbol{u}_1$, $\boldsymbol{u}_2$, …, $\boldsymbol{u}_n$ 是它的基底. 因 $\mathfrak{B}\cong\mathfrak{B}'$，故令这同构对应使 $\boldsymbol{u}_i\leftrightarrow\boldsymbol{u}_i'\in\mathfrak{B}'$ ($i=1, 2, \cdots, n$)，则 $\boldsymbol{u}_1'$, …, $\boldsymbol{u}_n'$ 是线性无关的：

这是因为，若 $k_1\boldsymbol{u}_1'+k_2\boldsymbol{u}_2'+\cdots+k_n\boldsymbol{u}_n'=\boldsymbol{0}$，则因从 $\mathfrak{B}\cong\mathfrak{B}'$ 中有 $\boldsymbol{u}_i\leftrightarrow\boldsymbol{u}_i'$ ($i=1, 2, \cdots, n$)，故有

$$k_1\boldsymbol{u}_1+k_2\boldsymbol{u}_2+\cdots+k_n\boldsymbol{u}_n\leftrightarrow k_1\boldsymbol{u}_1'+k_2\boldsymbol{u}_2'+\cdots+k_n\boldsymbol{u}_n'=\boldsymbol{0},$$

然同构对应 $\mathfrak{B}\cong\mathfrak{B}'$ 只能使 $\mathfrak{B}$ 之零向量对应于 $\mathfrak{B}'$ 之零向量，因此就有 $k_1\boldsymbol{u}_1+k_2\boldsymbol{u}_2+\cdots+k_n\boldsymbol{u}_n=\boldsymbol{0}$，于是由 $\boldsymbol{u}_1$, $\boldsymbol{u}_2$, …, $\boldsymbol{u}_n$ 线性无关，即得 $k_1=k_2=\cdots=k_n=0$，即证明了 $\boldsymbol{u}_1'$, …, $\boldsymbol{u}_n'$ 是线性无关的.

再若 $\boldsymbol{x}'\in\mathfrak{V}'$，而令 $\boldsymbol{x}=\lambda_1\boldsymbol{u}_1+\lambda_2\boldsymbol{u}_2+\cdots+\lambda_n\boldsymbol{u}_n$ 是 $\mathfrak{V}$ 中由 $\mathfrak{V}\cong\mathfrak{V}'$ 被 $\boldsymbol{x}'$ 所对应之向量，因为又有

$$\boldsymbol{x}=\lambda_1\boldsymbol{u}_1+\lambda_2\boldsymbol{u}_2+\cdots+\lambda_n\boldsymbol{u}_n\leftrightarrow\lambda_1\boldsymbol{u}_1'+\lambda_2\boldsymbol{u}_2'+\cdots+\lambda_n\boldsymbol{u}_n',$$

故必 $\boldsymbol{x}'=\lambda_1\boldsymbol{u}_1'+\lambda_2\boldsymbol{u}_2'+\cdots+\lambda_n\boldsymbol{u}_n'$. 这说明了 $\mathfrak{V}'$ 中任意向量都是 $\boldsymbol{u}_1'$, $\boldsymbol{u}_2'$, $\cdots$, $\boldsymbol{u}_n'$ 之线性组合，因而从它们之线性无关性知 $\boldsymbol{u}_1'$, $\cdots$, $\boldsymbol{u}_n'$ 为 $\mathfrak{V}'$ 之基底，于是 $\mathfrak{V}'$ 的维数 $=n$，证完.

练 习 十 四

1．设 $\mathfrak{V}$ 是变元 x 的一切小于 n 次多项式所形成的向量空间，试证：1，$(x-1)$，$(x-1)^2$，$\cdots$，$(x-1)^{n-1}$ 构成 $\mathfrak{V}$ 之一组基底；并求 $f(x)=a_0+a_1x+\cdots+a_{n-1}x^{n-1}$ 对于这基底的坐标.

2．假定 $\mathfrak{R}$ 是一切正实数之集合，依下面的规则进行运算：两个数的“加法”(记为 $\oplus$) 理解为它们通常的相乘 (即 $a\oplus b=ab$)，$\mathfrak{R}$ 中元 x 与实数 λ 之“乘法”(记为 $\odot$)理解为 x 之通常的 λ 次幂(即 $\lambda\odot x=x^\lambda$)．试证：$\mathfrak{R}$ 对于这两种运算形成一个 1 维的向量空间.

第四章　线性变换与矩阵代数

我们知道，在解析几何里，为了研究几何图形的变换，总是在一个固定的坐标系中进行讨论，以便把几何问题转化成为代数问题．同样，对 n 维空间几何学，为了便于将几何问题化为代数问题来讨论，也需要选定一组基底．本章就是讨论这个问题，详细研究线性变换及随之而来的矩阵运算．

第 1 节　平面和空间的旋转与投影变换

先谈平面的旋转变换和投影变换．如图 4-1，将平面上任一点 $P(x, y)$ 与原点 O 连接而成之向量 $\overrightarrow{OP}$，以反时针方向绕原点在平面上旋转 θ 角，得到向量 $\overrightarrow{OP'}$．即点 $P(x, y)$ 在平面上以反时针方向绕原点旋转 θ 角变到点 $P'(x', y')$．今研究这旋转变换的性质．为叙述简单，令 $\mathscr{T}$ 表示这旋转变换，记成

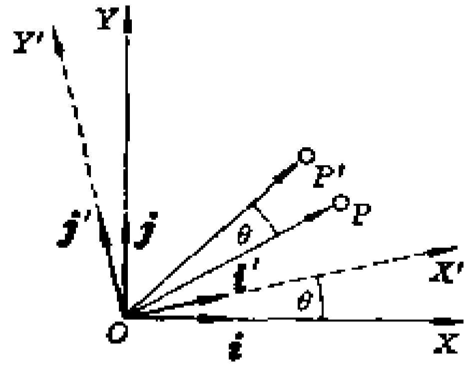


图　4-1

$$\overrightarrow{OP}\mathscr{T}=\overrightarrow{OP'}.$$

若将坐标系 XOY 也以反时针方向绕原点 O 旋转 θ 角，得新坐标系 $X'OY'$．旧坐标轴上基本单位向量 $\boldsymbol{i}, \boldsymbol{j}$ 变为新轴上的基本单位向量 $\boldsymbol{i}', \boldsymbol{j}'$，即 $\boldsymbol{i}\mathscr{T}=\boldsymbol{i}'$, $\boldsymbol{j}\mathscr{T}=\boldsymbol{j}'$，并易检验：假如 $\boldsymbol{\alpha}\mathscr{T}=\boldsymbol{\alpha}'$, $\boldsymbol{\beta}\mathscr{T}=\boldsymbol{\beta}'$，则必

$$(\boldsymbol{\alpha}+\boldsymbol{\beta})\mathscr{T}=\boldsymbol{\alpha}\mathscr{T}+\boldsymbol{\beta}\mathscr{T}=\boldsymbol{\alpha}'+\boldsymbol{\beta}',$$

$$(k\boldsymbol{\alpha})\mathscr{T}=k(\boldsymbol{\alpha}\mathscr{T})=k\boldsymbol{\alpha}'.$$

考察基本单位向量的变化，又易验证 $\boldsymbol{i}'$ 关于旧轴的坐标为 $\cos\theta$，$\sin\theta$，即

$$\boldsymbol{i}'=\boldsymbol{i}\cos\theta+\boldsymbol{j}\sin\theta.$$

同理，

$$\begin{aligned}\boldsymbol{j}'&=\boldsymbol{i}\cdot\cos\left(\frac{\pi}{2}+\theta\right)+\boldsymbol{j}\cdot\sin\left(\frac{\pi}{2}+\theta\right)\\&=-\boldsymbol{i}\cdot\sin\theta+\boldsymbol{j}\cdot\cos\theta.\end{aligned}$$

即

$$\begin{cases}\boldsymbol{i}\mathscr{T}=\boldsymbol{i}\cos\theta+\boldsymbol{j}\sin\theta,\\ \boldsymbol{j}\mathscr{T}=-\boldsymbol{i}\sin\theta+\boldsymbol{j}\cos\theta.\end{cases}\tag{1}$$

并将这式简单地记成

$$\begin{pmatrix}\boldsymbol{i}\\ \boldsymbol{j}\end{pmatrix}\mathscr{T}=\begin{pmatrix}\boldsymbol{i}\mathscr{T}\\ \boldsymbol{j}\mathscr{T}\end{pmatrix}=\begin{pmatrix}\boldsymbol{i}'\\ \boldsymbol{j}'\end{pmatrix}=\begin{pmatrix}\cos\theta & \sin\theta\\ -\sin\theta & \cos\theta\end{pmatrix}\cdot\begin{pmatrix}\boldsymbol{i}\\ \boldsymbol{j}\end{pmatrix},$$

式中，$\begin{pmatrix}\cos\theta & \sin\theta\\ -\sin\theta & \cos\theta\end{pmatrix}$称为向量之旋转变换 $\mathscr{T}$ 的矩阵.

容易求得 x, y 与 x', y' 的关系. 事实上，如令 $\boldsymbol{\alpha}=\overrightarrow{OP}$，则 $\boldsymbol{\alpha}=x\boldsymbol{i}+y\boldsymbol{j}$，因 P' 点关于新轴 $X'OY'$ 的相对位置同 P 点关于旧轴 XOY 的相对位置完全一样，故 $\boldsymbol{\alpha}\mathscr{T}=\boldsymbol{\alpha}'=x\boldsymbol{i}'+y\boldsymbol{j}'$；但又因为 $\boldsymbol{\alpha}'=x'\boldsymbol{i}+y'\boldsymbol{j}$，于是得到

$$\begin{aligned}x'\boldsymbol{i}+y'\boldsymbol{j}&=x\boldsymbol{i}'+y\boldsymbol{j}'\\&=x(\cos\theta\cdot\boldsymbol{i}+\sin\theta\cdot\boldsymbol{j})+y(-\sin\theta\cdot\boldsymbol{i}+\cos\theta\cdot\boldsymbol{j})\\&=(x\cos\theta-y\sin\theta)\boldsymbol{i}+(x\sin\theta+y\cos\theta)\boldsymbol{j},\end{aligned}$$

由 $\boldsymbol{i}$、$\boldsymbol{j}$ 之线性无关性，得知

$$\begin{cases}x'=x\cos\theta-y\sin\theta,\\ y'=x\sin\theta+y\cos\theta,\end{cases}$$

或简单地记成

$$\begin{pmatrix}x'\\ y'\end{pmatrix}=\begin{pmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{pmatrix}\begin{pmatrix}x\\ y\end{pmatrix},$$

式中，$\begin{pmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{pmatrix}$称为旋转变换$\mathscr{T}$诱导的向量坐标变换的矩阵，或简称点变换的矩阵．注意：这个点变换的矩阵

$$\begin{pmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{pmatrix}$$

和上述向量之旋转变换$\mathscr{T}$的矩阵

$$\begin{pmatrix}\cos\theta & \sin\theta\\ -\sin\theta & \cos\theta\end{pmatrix}$$

之差异是它们互为转置矩阵†.

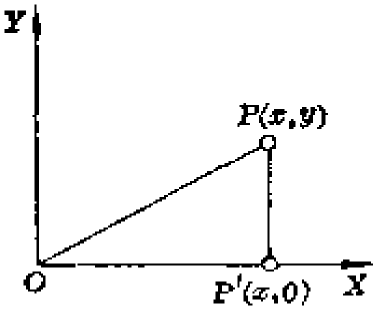


图 4-2

又，将平面上任一点$P(x,y)$与坐标系的原点O连接而得的向量$\overrightarrow{OP}$投射到x-轴上，得到向量$\overrightarrow{OP'}$(见图4-2)．于是，点$P(x, y)$变为点$P'(x, 0)$．如以$\mathscr{P}$表示这投影变换，即$\overrightarrow{OP}\mathscr{P}=\overrightarrow{OP'}$，则基本单位向量$\boldsymbol{i}, \boldsymbol{j}$分别变为$\boldsymbol{i}, \boldsymbol{0}$，即$\boldsymbol{i}\mathscr{P}=\boldsymbol{i}, \boldsymbol{j}\mathscr{P}=\boldsymbol{0}$，并易验证：若

$$\boldsymbol{\alpha}\mathscr{P}=\boldsymbol{\alpha}', \ \boldsymbol{\beta}\mathscr{P}=\boldsymbol{\beta}',$$

则

$$(\boldsymbol{\alpha}+\boldsymbol{\beta})\mathscr{P}=\boldsymbol{\alpha}\mathscr{P}+\boldsymbol{\beta}\mathscr{P}=\boldsymbol{\alpha}'+\boldsymbol{\beta}',$$

$$(k\boldsymbol{\alpha})\mathscr{P}=k(\boldsymbol{\alpha}\mathscr{P})=k\boldsymbol{\alpha}'.$$

这时，基本单位向量变化为

$$\begin{cases}\boldsymbol{i}\mathscr{P}=1\cdot\boldsymbol{i}+0\cdot\boldsymbol{j},\\ \boldsymbol{j}\mathscr{P}=0\cdot\boldsymbol{i}+0\cdot\boldsymbol{j},\end{cases} \tag{2}$$

或简写为

$$\begin{pmatrix}\boldsymbol{i}\\ \boldsymbol{j}\end{pmatrix}\mathscr{P}=\begin{pmatrix}\boldsymbol{i}\mathscr{P}\\ \boldsymbol{j}\mathscr{P}\end{pmatrix}=\begin{pmatrix}\boldsymbol{i}\\ \boldsymbol{0}\end{pmatrix}=\begin{pmatrix}1 & 0\\ 0 & 0\end{pmatrix}\begin{pmatrix}\boldsymbol{i}\\ \boldsymbol{j}\end{pmatrix},$$

† 对于$m\times n$-矩阵$\boldsymbol{A}=(a_{ij})$，定义$n\times m$-矩阵$\boldsymbol{B}=(b_{ij})$为$\boldsymbol{A}$的转置矩阵，其中$b_{ij}=a_{ji}$，并记成$\boldsymbol{B}=\boldsymbol{A}'$.

上式表明，向量的投影变换 $\mathscr{P}$（在 x-轴上）的矩阵为

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}.$$

又如令 $\boldsymbol{\alpha}=\overrightarrow{OP}$，则 $\boldsymbol{\alpha}=x\boldsymbol{i}+y\boldsymbol{j}$，$\boldsymbol{\alpha}\mathscr{P}=\boldsymbol{\alpha}'=x'\boldsymbol{i}+y'\boldsymbol{j}$，另方面，又有 $\boldsymbol{\alpha}\mathscr{P}=x\boldsymbol{i}$，故

$$\begin{cases} x'=1\cdot x+0\cdot y, \\ y'=0x+0y, \end{cases}$$

或简写为

$$\begin{pmatrix} x' \\ y' \end{pmatrix}=\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix},$$

这就是投影变换 $\mathscr{P}$ 在向量坐标间的变化公式，或点变换的变化公式，其矩阵 $\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ 仍可看做上述投影变换 $\mathscr{P}$ 的矩阵 $\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ 的转置矩阵．至于投影在 y-轴上的变换，可类似地讨论．

再考虑三维空间．如图 4-3，将空间中任一点 $P(x, y, z)$ 与坐标原点 O 连接而成的向量 $\overrightarrow{OP}$，绕着过原点 O 的某固定直线 l（以反时针方向）旋转 θ 角，而变到向量 $\overrightarrow{OP'}$；即点 $P(x, y, z)$ 变到点 $P'(x', y', z')$．如将这旋转变换记为 $\mathscr{T}$，即

$$\overrightarrow{OP}\mathscr{T}=\overrightarrow{OP'},$$

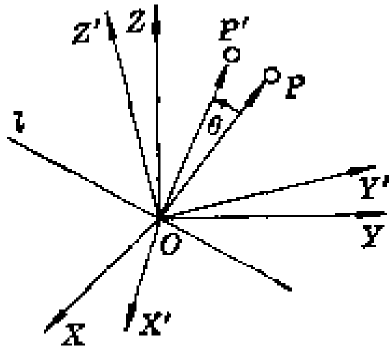


图 4-3

则与前述关于平面的情形类似，若将整个空间看做一个刚体而绕直线 l 旋转 θ 角之后，那末点 P 当然变到 P' 点，而旧坐标系 O-XYZ 变为新坐标系 O-$X'Y'Z'$，旧轴上的基本单位向量 $\boldsymbol{i}, \boldsymbol{j}, \boldsymbol{k}$ 就分别变为新轴上的基本单位向量 $\boldsymbol{i}', \boldsymbol{j}', \boldsymbol{k}'$，因之点 P' 关于新坐标系 O-$X'Y'Z'$

的位置关系恰如点 P 关于旧坐标系 $O\text{-}XYZ$ 的位置关系，故应有

$$\overrightarrow{OP'}=x'\boldsymbol{i}+y'\boldsymbol{j}+z'\boldsymbol{k}=x\boldsymbol{i}'+y\boldsymbol{j}'+z\boldsymbol{k}'.$$

若令 x', y', z' 轴关于坐标系 $O\text{-}XYZ$ 的方向余弦分别是 λ_1, μ_1, ν_1; λ_2, μ_2, ν_2; λ_3, μ_3, ν_3; 易知

$$\begin{cases}\boldsymbol{i}\mathscr{T}\,(=\boldsymbol{i}')=\lambda_1\boldsymbol{i}+\mu_1\boldsymbol{j}+\nu_1\boldsymbol{k},\\ \boldsymbol{j}\mathscr{T}\,(=\boldsymbol{j}')=\lambda_2\boldsymbol{i}+\mu_2\boldsymbol{j}+\nu_2\boldsymbol{k},\\ \boldsymbol{k}\mathscr{T}\,(=\boldsymbol{k}')=\lambda_3\boldsymbol{i}+\mu_3\boldsymbol{j}+\nu_3\boldsymbol{k},\end{cases}\tag{3}$$

或形式地简写为

$$\begin{pmatrix}\boldsymbol{i}\\ \boldsymbol{j}\\ \boldsymbol{k}\end{pmatrix}\mathscr{T}=\begin{pmatrix}\boldsymbol{i}\mathscr{T}\\ \boldsymbol{j}\mathscr{T}\\ \boldsymbol{k}\mathscr{T}\end{pmatrix}=\begin{pmatrix}\boldsymbol{i}'\\ \boldsymbol{j}'\\ \boldsymbol{k}'\end{pmatrix}=\begin{pmatrix}\lambda_1 & \mu_1 & \nu_1\\ \lambda_2 & \mu_2 & \nu_2\\ \lambda_3 & \mu_3 & \nu_3\end{pmatrix}\begin{pmatrix}\boldsymbol{i}\\ \boldsymbol{j}\\ \boldsymbol{k}\end{pmatrix},$$

式中，$\begin{pmatrix}\lambda_1 & \mu_1 & \nu_1\\ \lambda_2 & \mu_2 & \nu_2\\ \lambda_3 & \mu_3 & \nu_3\end{pmatrix}$称为向量之旋转变换 $\mathscr{T}$ 的矩阵.

如同平面解析几何中情况一样，这样的旋转变换 $\mathscr{T}$ 也具有下列性质：若 $\boldsymbol{\alpha}\mathscr{T}=\boldsymbol{\alpha}'$, $\boldsymbol{\beta}\mathscr{T}=\boldsymbol{\beta}'$, 则 $(\boldsymbol{\alpha}+\boldsymbol{\beta})\mathscr{T}=\boldsymbol{\alpha}\mathscr{T}+\boldsymbol{\beta}\mathscr{T}$; 且 $(k\boldsymbol{\alpha})\mathscr{T}=k(\boldsymbol{\alpha}\mathscr{T})$, 其中 k 为任意数.

下面来求 x', y', z' 与 x, y, z 的关系. 因

$$\begin{aligned}x'\boldsymbol{i}+y'\boldsymbol{j}+z'\boldsymbol{k}&=\overrightarrow{OP'}=x\boldsymbol{i}'+y\boldsymbol{j}'+z\boldsymbol{k}'\\ &=x(\lambda_1\boldsymbol{i}+\mu_1\boldsymbol{j}+\nu_1\boldsymbol{k})+y(\lambda_2\boldsymbol{i}+\mu_2\boldsymbol{j}+\nu_2\boldsymbol{k})\\ &\quad+z(\lambda_3\boldsymbol{i}+\mu_3\boldsymbol{j}+\nu_3\boldsymbol{k})\\ &=(\lambda_1x+\lambda_2y+\lambda_3z)\boldsymbol{i}+(\mu_1x+\mu_2y+\mu_3z)\boldsymbol{j}\\ &\quad+(\nu_1x+\nu_2y+\nu_3z)\boldsymbol{k},\end{aligned}$$

故有

$$\begin{cases}x'=\lambda_1x+\lambda_2y+\lambda_3z,\\ y'=\mu_1x+\mu_2y+\mu_3z,\\ z'=\nu_1x+\nu_2y+\nu_3z,\end{cases}$$

或形式地简写为

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix}=\begin{pmatrix} \lambda_1 & \lambda_2 & \lambda_3 \\ \mu_1 & \mu_2 & \mu_3 \\ \nu_1 & \nu_2 & \nu_3 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix},$$

式中，$\begin{pmatrix} \lambda_1 & \lambda_2 & \lambda_3 \\ \mu_1 & \mu_2 & \mu_3 \\ \nu_1 & \nu_2 & \nu_3 \end{pmatrix}$叫做由旋转变换$\mathscr{T}$所诱导的向量坐标变换的矩阵，简称点变换的矩阵．这矩阵显然为上述向量旋转变换的矩阵$\begin{pmatrix} \lambda_1 & \mu_1 & \nu_1 \\ \lambda_2 & \mu_2 & \nu_2 \\ \lambda_3 & \mu_3 & \nu_3 \end{pmatrix}$的转置矩阵．

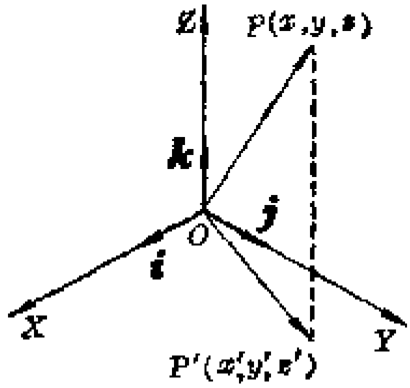


图 4-4

再如图 4-4，若将每向量

$$\overrightarrow{OP}=x\boldsymbol{i}+y\boldsymbol{j}+z\boldsymbol{k}$$

都投射到坐标面 XOY 上（至于投射到坐标面 YOZ、ZOX 上，得同样地处理），得到另一向量

$$\overrightarrow{OP'}=x'\boldsymbol{i}+y'\boldsymbol{j}+z'\boldsymbol{k},$$

即有 $x'=x$, $y'=y$, $z'=0$. 如把这投射变换表为$\mathscr{P}$，则也易验证：若 $\boldsymbol{\alpha}\mathscr{P}=\boldsymbol{\alpha}'$, $\boldsymbol{\beta}\mathscr{P}=\boldsymbol{\beta}'$，则

$$(\boldsymbol{\alpha}+\boldsymbol{\beta})\mathscr{P}=\boldsymbol{\alpha}\mathscr{P}+\boldsymbol{\beta}\mathscr{P}=\boldsymbol{\alpha}'+\boldsymbol{\beta}';$$

以及

$$(k\boldsymbol{\alpha})\mathscr{P}=k(\boldsymbol{\alpha}\mathscr{P})=k\boldsymbol{\alpha}'.$$

又因

$$\begin{cases} \boldsymbol{i}\mathscr{P}=\boldsymbol{i}=1\cdot\boldsymbol{i}+0\cdot\boldsymbol{j}+0\cdot\boldsymbol{k}, \\ \boldsymbol{j}\mathscr{P}=\boldsymbol{j}=0\cdot\boldsymbol{i}+1\cdot\boldsymbol{j}+0\cdot\boldsymbol{k}, \\ \boldsymbol{k}\mathscr{P}=\boldsymbol{0}=0\cdot\boldsymbol{i}+0\cdot\boldsymbol{j}+0\cdot\boldsymbol{k}, \end{cases} \tag{4}$$

或形式地简写为

$$\begin{pmatrix} \boldsymbol{i} \\ \boldsymbol{j} \\ \boldsymbol{k} \end{pmatrix} \mathscr{P} \left[= \begin{pmatrix} \boldsymbol{i}\mathscr{P} \\ \boldsymbol{j}\mathscr{P} \\ \boldsymbol{k}\mathscr{P} \end{pmatrix} = \begin{pmatrix} \boldsymbol{i} \\ \boldsymbol{j} \\ \boldsymbol{0} \end{pmatrix} \right] = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \boldsymbol{i} \\ \boldsymbol{j} \\ \boldsymbol{k} \end{pmatrix},$$

故这投影变换 $\mathscr{P}$ 的矩阵为

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

而将

$$\begin{cases} x' = x = 1\cdot x + 0\cdot y + 0\cdot z, \\ y' = y = 0\cdot x + 1\cdot y + 0\cdot z, \\ z' = 0 = 0\cdot x + 0\cdot y + 0\cdot z, \end{cases}$$

形式地表为

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

时，又得到了所谓点变换的矩阵 $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$，它仍然可以看做是向量的投影变换 $\mathscr{P}$ 的矩阵的转置矩阵.

第 2 节　n 维空间 $\mathfrak{R}_n$ 的线性变换

前节的讨论，完全属于直觉几何的范畴. 现在来讨论 n 维空间 $\mathfrak{R}_n$. 我们知道：象旋转和投影等这些具体概念，只是在直觉空间中(即 $n=2$, 3时)才有，对于 $n(\geqslant 4)$ 维空间 $\mathfrak{R}_n$ 怎样的向量变换叫做旋转变换或投影变换呢？ 在 $n(\geqslant 4)$ 维空间 $\mathfrak{R}_n$ 中，没有直觉的"旋转"或"投影"这些概念，只能象前章里

所说的那样，将直觉空间中($n=2$, 3 的情形)一些几何问题表成代数关系式，然后再将这些代数关系式推广到 $n\geqslant 4$ 的情形，而仍借用几何的术语来作解释.

我们知道：在平面与空间中，不论 $\mathscr{S}$ 是旋转变换或投影变换，$\mathscr{S}$ 具有下述性质：

$$(\boldsymbol{\alpha}+\boldsymbol{\beta})\mathscr{S}=\boldsymbol{\alpha}\mathscr{S}+\boldsymbol{\beta}\mathscr{S},\quad (k\boldsymbol{\alpha})\mathscr{S}=k(\boldsymbol{\alpha}\mathscr{S}).$$

于是，在 n 维空间 $\mathfrak{R}_n$ 中，我们就来研究具有这些性质的向量变换．在直觉空间中，向量变换的上述性质为旋转变换与投影变换所共有，因之只抓住这一性质的向量变换不能叫做旋转也不能叫做投影，注意到上述性质是线性函数的基本属性，于是就把它叫做线性变换，即

定义 如果有一个方法，能够使 n 维空间 $\mathfrak{R}_n$ 的每一个向量 $\boldsymbol{\alpha}$ 与 $\mathfrak{R}_n$ 中另一个由 $\boldsymbol{\alpha}$ 所完全确定的向量 $\boldsymbol{\alpha}'$ 对应，用符号 $\boldsymbol{\alpha}\to\boldsymbol{\alpha}'$ 来表示，这对应又叫做 $\mathfrak{R}_n$ 的变换，并叫 $\boldsymbol{\alpha}'$ 为 $\boldsymbol{\alpha}$ 的象，$\boldsymbol{\alpha}$ 为 $\boldsymbol{\alpha}'$ 的原象．如果这变换又适合下面的条件：

i) $\boldsymbol{\alpha}\to\boldsymbol{\alpha}'$, $\boldsymbol{\beta}\to\boldsymbol{\beta}'\Rightarrow\boldsymbol{\alpha}+\boldsymbol{\beta}\to\boldsymbol{\alpha}'+\boldsymbol{\beta}'$,

ii) $\boldsymbol{\alpha}\to\boldsymbol{\alpha}'\Rightarrow k\boldsymbol{\alpha}\to k\boldsymbol{\alpha}'$ (k为 数).

我们就叫这变换为 $\mathfrak{R}_n$ 的向量线性变换，简称 $\mathfrak{R}_n$ 的线性变换．这里，符号“$\Rightarrow$”表示“可导得”的意思.

如果用 $\mathscr{L}$ 表示线性变换：$\boldsymbol{\alpha}\to\boldsymbol{\alpha}'=\boldsymbol{\alpha}\mathscr{L}$，由定义知：对任意数 a, b，应恒有

$$(a\boldsymbol{\alpha}+b\boldsymbol{\beta})\mathscr{L}=a(\boldsymbol{\alpha}\mathscr{L})+b(\boldsymbol{\beta}\mathscr{L}).$$

反之，一变换 $\mathscr{L}$ 如具上面的性质，则 $\mathscr{L}$ 显为线性变换.

线性变换具有哪些基本性质呢？首先，它把一些特殊向量变成怎样的向量．当然，零向量是一个特殊向量，向量 $\boldsymbol{\alpha}$ 的负向量 $-\boldsymbol{\alpha}$ 也是一个特殊的(相对特殊的)向量．今设 $\mathscr{L}$ 为 $\mathfrak{R}_n$ 的线性变换：$\boldsymbol{\alpha}\to\boldsymbol{\alpha}\mathscr{L}=\boldsymbol{\alpha}'$．若 $\mathbf{0}\to\mathbf{0}\mathscr{L}=\mathbf{0}'$，则 $\boldsymbol{\alpha}=\boldsymbol{\alpha}+\mathbf{0}\to$

$(\boldsymbol{\alpha}+\mathbf{0})\mathscr{L}=\boldsymbol{\alpha}\mathscr{L}+\mathbf{0}\mathscr{L}=\boldsymbol{\alpha}'+\mathbf{0}'$. 故据对应之意义，应有 $\boldsymbol{\alpha}'+\mathbf{0}'=\boldsymbol{\alpha}'$，故 $\mathbf{0}'=\mathbf{0}$. 这说明了：在线性变换 $\mathscr{L}$ 下，零向量的象仍是零向量. 其次，若 $(-\boldsymbol{\alpha})\mathscr{L}=\boldsymbol{\gamma}$，则因

$$\mathbf{0}=\mathbf{0}\mathscr{L}=(\boldsymbol{\alpha}+(-\boldsymbol{\alpha}))\mathscr{L}=\boldsymbol{\alpha}\mathscr{L}+(-\boldsymbol{\alpha})\mathscr{L}=\boldsymbol{\alpha}'+\boldsymbol{\gamma},$$

故 $\boldsymbol{\gamma}=-\boldsymbol{\alpha}'=-(\boldsymbol{\alpha}\mathscr{L})$，这说明了：在线性变换 $\mathscr{L}$ 下，$\boldsymbol{\alpha}$ 的负向量 $-\boldsymbol{\alpha}$ 的象等于 $\boldsymbol{\alpha}$ 之象的负. 于是，有

性质 1 $\mathfrak{R}_n$ 之线性变换必使 $\mathbf{0}$ 变为 $\mathbf{0}$，使 $\boldsymbol{\alpha}$ 之负（即 $-\boldsymbol{\alpha}$）变为 $\boldsymbol{\alpha}$ 之象的负.

注意：所谓 $\mathscr{L}$ 是 $\mathfrak{R}_n$ 之变换，指的是对于每 $\boldsymbol{\alpha}(\in\mathfrak{R}_n)$ 都有一个完全确定的象 $\boldsymbol{\alpha}'(=\boldsymbol{\alpha}\mathscr{L})$，但两个不同的向量可能有相同的象，即虽 $\boldsymbol{\alpha}\neq\boldsymbol{\beta}$，而有 $\boldsymbol{\alpha}\mathscr{L}=\boldsymbol{\beta}\mathscr{L}$. 再，$\mathfrak{R}_n$ 的每个向量不一定都有原象，即虽 $\boldsymbol{\alpha}'\in\mathfrak{R}_n$，但不一定有 $\boldsymbol{\alpha}\in\mathfrak{R}_n$，使 $\boldsymbol{\alpha}\mathscr{L}=\boldsymbol{\alpha}'$. 例如，使 $\mathfrak{R}_n$ 的 $\boldsymbol{\alpha}=(a_1, a_2, \cdots, a_n)$ 对应于 $\boldsymbol{\alpha}'=(a_1, 0, \cdots, 0)$ 的变换 $\mathscr{T}$，易证 $\mathscr{T}$ 是一个线性变换，而且虽 $\boldsymbol{\alpha}=(a_1, a_2, \cdots, a_n)\neq\boldsymbol{\beta}=(a_1, b_2, \cdots, b_n)$，但确有 $\boldsymbol{\alpha}'=(a_1, 0, \cdots, 0)=\boldsymbol{\beta}'$，即 $\boldsymbol{\alpha}\mathscr{T}=\boldsymbol{\beta}\mathscr{T}$；对于这样的 $\mathscr{T}$，不是 $\mathfrak{R}_n$ 之每个向量都有原象，例如向量 $(1, 1, 0, \cdots, 0)$ 就没有原象.

对于 $\mathfrak{R}_n$ 之线性变换 $\mathscr{T}$，如果 $\mathfrak{R}_n$ 中有一向量不止一个原象，那末 $\mathfrak{R}_n$ 之 $\mathbf{0}$ 的原象也就不只是 $\mathbf{0}$†；反之，若 $\mathbf{0}$ 之原象不只是 $\mathbf{0}$ 时，则 $\mathfrak{R}_n$ 之向量当它有原象时都必有一个以上的原象††. 这就证得了

性质 2 对于 $\mathfrak{R}_n$ 之线性变换 $\mathscr{T}$，使零向量 $\mathbf{0}$ 之原象只能是 $\mathbf{0}$ 的充要条件是：每 $\boldsymbol{\alpha}(\in\mathfrak{R}_n)$ 之原象如存在，则必为唯一的.

---

† 例如 $\boldsymbol{\alpha}\mathscr{T}=\boldsymbol{\beta}\mathscr{T}=\boldsymbol{\alpha}'(\boldsymbol{\alpha}\neq\boldsymbol{\beta})\Rightarrow(\boldsymbol{\alpha}-\boldsymbol{\beta})\mathscr{T}=\boldsymbol{\alpha}\mathscr{T}-\boldsymbol{\beta}\mathscr{T}=\mathbf{0}$ 且 $\boldsymbol{\alpha}-\boldsymbol{\beta}\neq\mathbf{0}$.

†† 事实上，设 $\boldsymbol{\alpha}\mathscr{T}=\mathbf{0}$，$\boldsymbol{\alpha}\neq\mathbf{0}$，则从 $\boldsymbol{\beta}\mathscr{T}=\boldsymbol{\beta}'$ 即得 $(\boldsymbol{\alpha}+\boldsymbol{\beta})\mathscr{T}=\boldsymbol{\alpha}\mathscr{T}+\boldsymbol{\beta}\mathscr{T}=\mathbf{0}+\boldsymbol{\beta}'=\boldsymbol{\beta}'$，并有 $\boldsymbol{\alpha}+\boldsymbol{\beta}\neq\boldsymbol{\beta}$.

今问：对线性变换 $\mathscr{L}$，$\mathfrak{R}_n$ 之 $\mathbf{0}$ 的原象不只是零向量时，那末 $\mathbf{0}$ 之一切原象的集合 $\mathfrak{M}$ 究竟是什么？

若 $\boldsymbol{\alpha}$，$\boldsymbol{\beta}\in\mathfrak{M}$，则 $\boldsymbol{\alpha}\mathscr{L}=\mathbf{0}=\boldsymbol{\beta}\mathscr{L}$，因此 $(\boldsymbol{\alpha}+\boldsymbol{\beta})\mathscr{L}=\boldsymbol{\alpha}\mathscr{L}+\boldsymbol{\beta}\mathscr{L}=\mathbf{0}$ 与 $(k\boldsymbol{\alpha})\mathscr{L}=k(\boldsymbol{\alpha}\mathscr{L})=\mathbf{0}$，说明了 $\boldsymbol{\alpha}+\boldsymbol{\beta}\in\mathfrak{M}$ 及 $k\boldsymbol{\alpha}\in\mathfrak{M}$，即 $\mathfrak{M}$ 为 $\mathfrak{R}_n$ 之子空间．故得

性质 3 对于 $\mathfrak{R}_n$ 的任意线性变换 $\mathscr{L}$，$\mathfrak{R}_n$ 中零元的一切原象的集合 $\mathfrak{M}$，形成 $\mathfrak{R}_n$ 之子空间．

上面已说过：对于某线性变换 $\mathscr{L}$，$\mathfrak{R}_n$ 之每个向量不一定都有原象．因此，线性变换 $\mathscr{L}$ 将整个空间 $\mathfrak{R}_n$ 可能变为它的真子集合 $\mathfrak{N}$，那么子集合 $\mathfrak{N}$ 有何性质呢？

若 $\boldsymbol{\alpha}'\in\mathfrak{N}$，则 $\boldsymbol{\alpha}'$ 至少有一原象，令 $\boldsymbol{\alpha}$ 为 $\boldsymbol{\alpha}'$ 的一个原象，即 $\boldsymbol{\alpha}'=\boldsymbol{\alpha}\mathscr{L}$．因对任意数 k，有 $(k\boldsymbol{\alpha})\mathscr{L}=k(\boldsymbol{\alpha}\mathscr{L})=k\boldsymbol{\alpha}'$，故 $k\boldsymbol{\alpha}'\in\mathfrak{N}$；其次，从 $\boldsymbol{\alpha}'\in\mathfrak{N}$，$\boldsymbol{\beta}'\in\mathfrak{N}$，知有 $\boldsymbol{\alpha}$、$\boldsymbol{\beta}$ 使 $\boldsymbol{\alpha}\mathscr{L}=\boldsymbol{\alpha}'$，$\boldsymbol{\beta}\mathscr{L}=\boldsymbol{\beta}'$，于是 $(\boldsymbol{\alpha}+\boldsymbol{\beta})\mathscr{L}=\boldsymbol{\alpha}\mathscr{L}+\boldsymbol{\beta}\mathscr{L}=\boldsymbol{\alpha}'+\boldsymbol{\beta}'$，就说明了 $\boldsymbol{\alpha}'+\boldsymbol{\beta}'\in\mathfrak{N}$．故 $\mathfrak{N}$ 为 $\mathfrak{R}_n$ 的子空间，即证得了：

性质 4 设 $\mathscr{L}$ 为 $\mathfrak{R}_n$ 的一线性变换，$\mathfrak{N}$ 是 $\mathfrak{R}_n$ 中所有向量的象之集合，记为 $\mathfrak{N}=\mathfrak{R}_n\mathscr{L}$，则 $\mathfrak{N}$ 为 $\mathfrak{R}_n$ 之子空间．

我们知道：线性变换这个概念是直觉空间中旋转与投影这两个概念的抽象，研究起来带有一定的困难，如何使线性变换这概念具体化，即用怎样的方式来具体地表现它呢．

回顾前节，有所谓向量的旋转变换和投影变换的矩阵．然而产生矩阵要以坐标系中的基本单位向量为背景，可参看(1)，(2)，(3)，(4)诸式．但在 n 维空间 $\mathfrak{R}_n$ 内，相当于直觉空间中基本单位向量的推广是基底，于是我们就着手研究 $\mathfrak{R}_n$ 的线性变换对于 $\mathfrak{R}_n$ 的一组基底所起的作用．

设 $\mathscr{T}$ 为 $\mathfrak{R}_n$ 之一线性变换，取 $\mathfrak{R}_n$ 之一组基底 $\boldsymbol{\alpha}_1$，$\boldsymbol{\alpha}_2$，…，$\boldsymbol{\alpha}_n$，因而每个 $\boldsymbol{\alpha}_i\mathscr{T}$ 得写为 $\boldsymbol{\alpha}_1$，$\boldsymbol{\alpha}_2$，…，$\boldsymbol{\alpha}_n$ 的线性组合，如

$$\begin{cases} \boldsymbol{\alpha}_1\mathscr{T}=a_{11}\boldsymbol{\alpha}_1+a_{12}\boldsymbol{\alpha}_2+\cdots+a_{1n}\boldsymbol{\alpha}_n, \\ \boldsymbol{\alpha}_2\mathscr{T}=a_{21}\boldsymbol{\alpha}_1+a_{22}\boldsymbol{\alpha}_2+\cdots+a_{2n}\boldsymbol{\alpha}_n, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ \boldsymbol{\alpha}_n\mathscr{T}=a_{n1}\boldsymbol{\alpha}_1+a_{n2}\boldsymbol{\alpha}_2+\cdots+a_{nn}\boldsymbol{\alpha}_n. \end{cases} \tag{5}$$

由(5)得一个 n 级矩阵

$$\boldsymbol{A}=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}.$$

这就是说：关于给定的一组基底 $\boldsymbol{\alpha}_1, \cdots, \boldsymbol{\alpha}_n$，向量空间 $\mathfrak{R}_n$ 的每个线性变换 $\mathscr{T}$ 都可通过(5)式对应唯一个矩阵 $\boldsymbol{A}$. 这样，就可以设想把线性变换的研究转化为矩阵的研究. 但是，要实现这种转化的可能性，必须先解决下面两个问题.

第一个问题是：若事先给了一 n 级矩阵 $\boldsymbol{A}$，是否反过来也必有一个能满足(5)式的线性变换 $\mathscr{T}$ 呢？如果这样的 $\mathscr{T}$ 存在，才有可能所谓把线性变换转化为矩阵，即具体的一个矩阵确实可以表示一个抽象的线性变换.

第二个问题是：即令第一个问题解决了，还怕遇见两个不同的线性变换通过(5)式对应于同一个矩阵(若是这样，将线性变换转化为矩阵时就会发生混淆，因为如果一个矩阵可能表示两个不同的线性变换，所以对这个矩阵的研究究竟反映了哪一个变换的性质呢?)，故还必须解决一个问题：一个矩阵只能被唯一个线性变换所对应.

由矩阵 $\boldsymbol{A}=(a_{ij})$ 的任意性，得知(5)式右边的 n 个向量 $\sum\limits_{i=1}^{n}a_{1i}\boldsymbol{\alpha}_i, \sum\limits_{i=1}^{n}a_{2i}\boldsymbol{\alpha}_i, \cdots, \sum\limits_{i=1}^{n}a_{ni}\boldsymbol{\alpha}_i$ 是任意的，若分别用 $\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \cdots, \boldsymbol{\beta}_n$ 来表示，则上述第一个问题就是要证明必有一线性变换 $\mathscr{T}$ 使 $\boldsymbol{\alpha}_i\mathscr{T}=\boldsymbol{\beta}_i(i=1, 2, \cdots, n)$，而第二个问题则是要证明这样

的 $\mathscr{T}$ 还是唯一的．下面的基本定理就肯定了这个答案．

基本定理 设 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 是 $\Re_n$ 之基底，而 $\beta_1, \beta_2, \cdots, \beta_n$ 是 $\Re_n$ 之任 n 个向量，则有一个且仅一个线性变换 $\mathscr{T}$，使 $\alpha_i\mathscr{T}=\beta_i(i=1, 2, \cdots, n)$．

证明 因 $\Re_n$ 之任意向量 α 可唯一地表为

$$\alpha=k_1\alpha_1+k_2\alpha_2+\cdots+k_n\alpha_n,$$

故可作出一个变换 $\mathscr{T}$，使 $\alpha\mathscr{T}=k_1\beta_1+k_2\beta_2+\cdots+k_n\beta_n$（即把 α 变换成 $\alpha\mathscr{T}$）．显然，据 $\mathscr{T}$ 之形成的意义，应有 $\alpha_i\mathscr{T}=\beta_i(i=1, 2, \cdots, n)$．下面来证明 $\mathscr{T}$ 为 $\Re_n$ 之线性变换：这是因为对任意数 k，有

$$\begin{aligned} k\alpha &= k(k_1\alpha_1+k_2\alpha_2+\cdots+k_n\alpha_n) \\ &= k(k_1\alpha_1)+k(k_2\alpha_2)+\cdots+k(k_n\alpha_n) \\ &= \sum_{i=1}^{n}(kk_i)\alpha_i, \end{aligned}$$

故由形成 $\mathscr{T}$ 之意义，得

$$\begin{aligned} (k\alpha)\mathscr{T} &= \sum_{i=1}^{n}(kk_i)\beta_i=\sum_{i=1}^{n}k(k_i\beta_i) \\ &= k\cdot\sum_{i=1}^{n}k_i\beta_i=k\cdot(\alpha\mathscr{T}); \end{aligned}$$

其次，若 $\gamma=\sum_{i=1}^{n}l_i\alpha_i$，则又有 $\gamma\mathscr{T}=\sum_{i=1}^{n}l_i\beta_i$，因而从

$$\alpha+\gamma=\sum_{i=1}^{n}(k_i+l_i)\alpha_i$$

又得

$$\begin{aligned} (\alpha+\gamma)\mathscr{T} &= \sum_{i=1}^{n}(k_i+l_i)\beta_i=\sum_{i=1}^{n}k_i\beta_i+\sum_{i=1}^{n}l_i\beta_i \\ &= \alpha\mathscr{T}+\gamma\mathscr{T}. \end{aligned}$$

这就证得了 $\mathscr{T}$ 确为 $\Re_n$ 之线性变换．这样，就证得了定理中所说的 $\mathscr{T}$（线性变换）之存在性．

再证明线性变换 $\mathscr{T}$ 之唯一性. 设有 $\mathfrak{R}_n$ 之线性变换 $\mathscr{T}$ 和 $\mathscr{L}$, 均使 $\boldsymbol{\alpha}_i$ 变为 $\boldsymbol{\beta}_i(i=1, 2, \cdots, n)$, 即 $\boldsymbol{\alpha}_i\mathscr{T}=\boldsymbol{\beta}_i$, $\boldsymbol{\alpha}_i\mathscr{L}=\boldsymbol{\beta}_i$, 根据线性变换的意义,有

$$\boldsymbol{\alpha}\mathscr{L}=\left(\sum_{i=1}^{n}k_i\boldsymbol{\alpha}_i\right)\mathscr{L}=\sum_{i=1}^{n}(k_i\boldsymbol{\alpha}_i)\mathscr{L}$$

$$=\sum_{i=1}^{n}k_i(\boldsymbol{\alpha}_i\mathscr{L})=\sum_{i=1}^{n}k_i\boldsymbol{\beta}_i=\boldsymbol{\alpha}\mathscr{T}.$$

这说明了两个线性变换 $\mathscr{L}$ 与 $\mathscr{T}$ 作用在 $\mathfrak{R}_n$ 上是一样的,故必 $\mathscr{L}=\mathscr{T}$.

基本定理因而获证.

根据基本定理, 前述的第一、二两个问题都解决了. 因之, 对于基底 $\boldsymbol{\alpha}_i(i=1, 2, \cdots, n)$, 得知 $\mathfrak{R}_n$ 之一切线性变换与一切 n 级矩阵之间存在着一一对应关系. 正因为这样, 有时又常用符号 $\mathscr{T}_A$ 表示矩阵为 $\boldsymbol{A}$ 的线性变换. 于是(5)形式地可写为

$$\begin{pmatrix}\boldsymbol{\alpha}_1\\ \boldsymbol{\alpha}_2\\ \vdots\\ \boldsymbol{\alpha}_n\end{pmatrix}\mathscr{T}_A=\begin{pmatrix}\boldsymbol{\alpha}_1\mathscr{T}_A\\ \boldsymbol{\alpha}_2\mathscr{T}_A\\ \vdots\\ \boldsymbol{\alpha}_n\mathscr{T}_A\end{pmatrix}=\boldsymbol{A}\begin{pmatrix}\boldsymbol{\alpha}_1\\ \boldsymbol{\alpha}_2\\ \vdots\\ \boldsymbol{\alpha}_n\end{pmatrix},$$

这里,矩阵 $\boldsymbol{A}$ 称为线性变换 $\mathscr{T}_A$ 的矩阵(对基底 $\boldsymbol{\alpha}_1, \cdots, \boldsymbol{\alpha}_n$), 它的行列式

$$D_A=\begin{vmatrix}a_{11} & a_{12} & \cdots & a_{1n}\\ a_{21} & a_{22} & \cdots & a_{2n}\\ \vdots & \vdots & & \vdots\\ a_{n1} & a_{n2} & \cdots & a_{nn}\end{vmatrix}$$

叫做 $\mathscr{T}_A$ 的行列式. 如果 $D_A\neq0$, $\mathscr{T}_A$ 和 $\boldsymbol{A}$ 都叫满秩的(或非奇异的); 若 $D_A=0$, 则叫 $\mathscr{T}_A$ 与 $\boldsymbol{A}$ 都是降秩的(或奇异的).

基本定理解决了这样一个重要问题，即抽象的线性变换可用具体的矩阵来表示，也就是肯定了它们之间存在着一一对应关系．因之，几何问题（线性变换）便可以转化为代数问题（矩阵）来研究了．由于这种转化的依据是上述定理，这正是我们把这个定理称为“基本定理”的原因．

这里，有几点要提请读者注意：

第一，向量 $\boldsymbol{\alpha}=x_1\boldsymbol{\alpha}_1+x_2\boldsymbol{\alpha}_2+\cdots+x_n\boldsymbol{\alpha}_n$ 关于基底 $\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_n$ 之坐标 $x_1, x_2, \cdots, x_n$，经过线性变换 $\mathscr{T}_A$ 后会发生怎样的变化呢？

令 $\boldsymbol{\alpha}\mathscr{T}_A=y_1\boldsymbol{\alpha}_1+y_2\boldsymbol{\alpha}_2+\cdots+y_n\boldsymbol{\alpha}_n$．于是利用(5)式易知

$$\boldsymbol{\alpha}\mathscr{T}_A=\left(\sum_{i=1}^n x_i\boldsymbol{\alpha}_i\right)\mathscr{T}_A=\sum_{i=1}^n x_i(\boldsymbol{\alpha}_i\mathscr{T}_A)$$

$$=\sum_{i=1}^n x_i\left(\sum_{j=1}^n a_{ij}\boldsymbol{\alpha}_j\right)=\sum_{j=1}^n\left(\sum_{i=1}^n a_{ij}x_i\right)\boldsymbol{\alpha}_j,$$

故有

$$\begin{cases} y_1=a_{11}x_1+a_{21}x_2+\cdots+a_{n1}x_n, \\ y_2=a_{12}x_1+a_{22}x_2+\cdots+a_{n2}x_n, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ y_n=a_{1n}x_1+a_{2n}x_2+\cdots+a_{nn}x_n, \end{cases} \tag{6}$$

这就是由 $\mathscr{T}_A$ 产生的把变数 $x_1, \cdots, x_n$ 换成 $y_1, \cdots, y_n$ 的线性变换，即产生 $\mathfrak{R}_n$ 中的点的线性变换（也就是把原向量 $\boldsymbol{\alpha}$ 的坐标换成它的象 $\boldsymbol{\alpha}\mathscr{T}_A$ 的坐标的线性变换）．并且，这个点线性变换的矩阵是 $\boldsymbol{A}$ 之转置矩阵，这就是空间向量的线性变换和点的线性变换的关系，这和在第1节所说的直觉空间中旋转与投影的情况是完全一致的．

第二，所谓线性变换与矩阵（n 级）间存在一一对应关系，是指对 $\mathfrak{R}_n$ 的一组基底 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_n$ 来说的；如果对另一组基底 $\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \cdots, \boldsymbol{\beta}_n$，结果又怎样呢？

对于基底 $\boldsymbol{\alpha}_1, \cdots, \boldsymbol{\alpha}_n$, 设线性变换 $\mathscr{T}$ 之矩阵为 $\boldsymbol{A}=(a_{ij})$, 即有 $\boldsymbol{\alpha}_i\mathscr{T}=\sum_{j=1}^{n} a_{ij}\boldsymbol{\alpha}_j$, 而对于基底 $\boldsymbol{\beta}_1, \cdots, \boldsymbol{\beta}_n$, $\mathscr{T}$ 之矩阵为 $\boldsymbol{B}=(b_{ij})$, 同样有 $\boldsymbol{\beta}_i\mathscr{T}=\sum_{j=1}^{n} b_{ij}\boldsymbol{\beta}_j$. 由于一般来讲不一定有 $\boldsymbol{A}=\boldsymbol{B}$, 这就是说: $\mathfrak{R}_n$ 之基底不同, 同一个线性变换 $\mathscr{T}$ 的矩阵一般也不同. 它们之间有怎样的关系呢? 这将在第 6 节里再详细地讨论.

第三, 由 $\mathfrak{R}_n$ 之线性变换 $\mathscr{T}_A$ 如性质 3 及 4 说的所产生的两个子空间 $\mathfrak{M}$ 及 $\mathfrak{N}$ 有什么关系呢?

设矩阵 $\boldsymbol{A}$ 的秩为 r. 于是 $\boldsymbol{A}$ 中线性无关的行向量之最大数目为 r, 故这时(5)式里 n 个向量 $\boldsymbol{\alpha}_1\mathscr{T}_A, \cdots, \boldsymbol{\alpha}_n\mathscr{T}_A$ 中线性无关的最大个数为 r. 由于 $\mathfrak{N}=\mathfrak{R}_n\mathscr{T}_A$ 中每个向量为 $\boldsymbol{\alpha}_1\mathscr{T}_A, \cdots, \boldsymbol{\alpha}_n\mathscr{T}_A$ 的线性组合, 故 $\mathfrak{N}$ 之维数为 r, 即 $\mathfrak{N}$ 的维数等于矩阵 $\boldsymbol{A}$ 之秩. 从而可知, $\mathfrak{N}=\mathfrak{R}_n$ 的充要条件是 $\boldsymbol{A}$ 为满秩的.

再来研究子空间 $\mathfrak{M}$. $\boldsymbol{x}=\sum_{i=1}^{n} c_i\boldsymbol{\alpha}_i\in\mathfrak{M}$ 的充要条件是

$$\begin{aligned}\mathbf{0}=\boldsymbol{x}\mathscr{T}_A&=\sum_{i=1}^{n} c_i(\boldsymbol{\alpha}_i\mathscr{T}_A)=\sum_{i=1}^{n} c_i\sum_{j=1}^{n} a_{ij}\boldsymbol{\alpha}_j\\&=\sum_{j=1}^{n}\left(\sum_{i=1}^{n} a_{ij}c_i\right)\boldsymbol{\alpha}_j,\end{aligned}$$

随而又为

$$\sum_{i=1}^{n} a_{ij}c_i=0 \quad (j=1, 2, \cdots, n). \tag{7}$$

因 $\boldsymbol{A}$ 之秩为 r, 故(7)式有 $n-r$ 个线性无关的解:

$$\boldsymbol{\gamma}_1=(c_1^{(1)}, \cdots, c_n^{(1)}), \cdots, \boldsymbol{\gamma}_{n-r}=(c_1^{(n-r)}, \cdots, c_n^{(n-r)}),$$

使(7)之任何解 $\boldsymbol{\gamma}=(c_1, \cdots, c_n)$ 为 $\boldsymbol{\gamma}_1, \cdots, \boldsymbol{\gamma}_{n-r}$ 的线性组合, 即

$$\boldsymbol{\gamma}=\lambda_1\boldsymbol{\gamma}_1+\cdots+\lambda_{n-r}\boldsymbol{\gamma}_{n-r}. \tag{8}$$

故 $\boldsymbol{x}=\sum_{i=1}^{n}c_i\boldsymbol{\alpha}_i\in\mathfrak{M}$ 的充要条件是存在$\lambda_1,\cdots,\lambda_{n-r}$使(8)成立.但因 $\boldsymbol{x}_i=c_1^{(i)}\boldsymbol{\alpha}_1+\cdots+c_n^{(i)}\boldsymbol{\alpha}_n(i=1,2,\cdots,n-r)$为 $\mathfrak{M}$ 中的$n-r$个线性无关的向量,故 $\boldsymbol{x}\in\mathfrak{M}$ 的充要条件是

$$\boldsymbol{x}=\lambda_1\boldsymbol{x}_1+\lambda_2\boldsymbol{x}_2+\cdots+\lambda_{n-r}\boldsymbol{x}_{n-r},$$

即 $\boldsymbol{x}$ 得表为 $\boldsymbol{x}_1,\cdots,\boldsymbol{x}_{n-r}$ 的线性组合. 这证明了 $\mathfrak{M}$ 的维数等于$n-r$. 即证得了:

定理 1 设$\mathscr{T}$是空间 $\mathfrak{R}_n$ 的线性变换,则 $\mathfrak{R}_n$ 中所有各元的象 $\mathfrak{N}=\mathfrak{R}_n\mathscr{T}$ 构成 $\mathfrak{R}_n$ 的子空间之维数以及 $\mathfrak{R}_n$ 之零元的一切原象之集 $\mathfrak{M}$ 而成之子空间的维数,两者的和等于n. 而且,不论 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$ 是 $\mathfrak{R}_n$ 之任何一组基底,$\mathscr{T}$ 关于这组基底之矩阵的秩就等于 $\mathfrak{N}$ 的维数, 因而 $\mathscr{T}$ 所对应之矩阵的秩为一常数(即与 $\mathscr{T}$ 自身有关而与基底之选择无关).

注意: 虽然定理 1 说了 $\mathfrak{N}$ 之维数$+\mathfrak{M}$ 之维数$=n$($\mathfrak{R}_n$ 的维数),但 $\mathfrak{N}\cup\mathfrak{M}$ 并不是全空间 $\mathfrak{R}_n$,即可能有 $\mathfrak{N}\cup\mathfrak{M}<\mathfrak{R}_n$ 之情况. 例如,令 $\mathfrak{P}[x]_n$ 表示变量x的所有小于n次多项式之集而成的n维向量空间, 则 $\mathscr{D}(f(x))=f'(x)$, 即微商变换 $\mathscr{D}$,显然是 $\mathfrak{P}[x]_n$ 的一个线性变换, 而 $\mathscr{D}(\mathfrak{P}[x]_n)=\mathfrak{N}$ 在此时显为 $\mathfrak{P}[x]_{n-1}$,即 $\mathfrak{N}=\mathfrak{P}[x]_{n-1}$,又使 $\mathscr{D}(f(x))=0$ 之一切$f(x)$而成之子空间 $\mathfrak{M}$ 在此时显为数域 $\mathfrak{P}$,即 $\mathfrak{M}=\mathfrak{P}$, 故

$$\mathfrak{M}\cup\mathfrak{N}=\mathfrak{P}\cup\mathfrak{P}[x]_{n-1}=\mathfrak{P}[x]_{n-1}<\mathfrak{P}[x]_n.$$

由定理 1, 不难得知:

推论 对于线性变换 $\mathscr{T}$,空间 $\mathfrak{R}_n$ 之一向量如有原象,其原象仅有一个的充要条件是 $\mathfrak{R}_n$ 的每个向量都有原象.

证明 如有原象时,由前述关于线性变换的性质 2, 原象只一个的充要条件是零向量的原象只能为 $\mathbf{0}$, 亦即 $\mathfrak{M}=\mathfrak{O}$, 因之 $\mathfrak{N}=\mathfrak{R}_n$, 此即说明每个向量都有原象.

练 习 十 五

1．设 α_1, α_2, α_3, α_4 为 4 维向量空间 $\mathfrak{R}_4$ 的基底，线性变换 $\mathscr{L}$ 对于这组基底的矩阵是

$$\begin{pmatrix} 1 & 2 & 3 & 2 \\ -1 & 0 & 3 & 1 \\ 2 & 1 & 5 & -1 \\ 1 & 1 & 2 & 2 \end{pmatrix}.$$

证明 $\beta_1(=\alpha_1)$, $\beta_2(=\alpha_1+\alpha_2)$, $\beta_3(=\alpha_1+\alpha_2+\alpha_3)$, $\beta_4(=\alpha_1+\alpha_2+\alpha_3+\alpha_4)$ 亦为 $\mathfrak{R}_4$ 之基底，并求 $\mathscr{L}$ 对这组基底的矩阵．

2．下面的各个变换 $\mathscr{L}$ 中，哪些是线性的，哪些不是：

1) 在 $\mathfrak{R}_n$ 中，$\alpha\mathscr{L}=\alpha+\gamma$，这里 $\gamma\in\mathfrak{R}_n$ 是一个固定的向量；

2) 在 $\mathfrak{R}_n$ 中，$\alpha\mathscr{L}=\gamma$，这里 $\gamma\in\mathfrak{R}_n$ 是固定的；

3) 设 $\mathfrak{C}$ 为复数域，并视 $\mathfrak{C}$ 为 $\mathfrak{C}$ 上的线性空间，$\alpha\mathscr{L}=\bar{\alpha}$；

4) 若将 3) 中 $\mathfrak{C}$ 视为 $\mathfrak{R}^*$ 上的线性空间（$\mathfrak{R}^*$ 为实数域），$\alpha\mathscr{L}=\bar{\alpha}$；

5) 在 $\mathfrak{R}_3$ 中，$(x_1, x_2, x_3)\mathscr{L}=(x_1^2, x_2, x_3)$．

3．在 $\mathfrak{R}_2$ 中，令每个向量都投影到直线 $ax+by=0$ 上 $(ab\neq0)$，证明这投影变换 $\mathscr{T}$ 是 $\mathfrak{R}_2$ 的线性变换，并求它对于基底 $\boldsymbol{i}$, $\boldsymbol{j}$ 的矩阵．

第 3 节　线性变换和矩阵的运算

这里，介绍向量空间 $\mathfrak{R}_n$ 中线性变换之间的种种运算，从而定义 n 级矩阵之间的种种运算．

在这节里，恒假定 $\mathfrak{R}_n$ 的一组基底固定为 α_1, α_2, $\cdots$, α_n．因此，$\mathfrak{R}_n$ 中线性变换的矩阵总是指的关于这组固定的基底而言的，后而不另声明．

1) 线性变换的加法　假设给了 $\mathfrak{R}_n$ 之两个线性变换 $\mathscr{T}_A$ 与 $\mathscr{T}_B$．令 $\mathscr{T}$ 表示 $\mathfrak{R}_n$ 的一个这样的变换，即使每 $\alpha\in\mathfrak{R}_n$ 对应于 $\alpha\mathscr{T}=\alpha\mathscr{T}_A+\alpha\mathscr{T}_B$．

由于

$$
\begin{aligned}
(k\boldsymbol{\alpha}+l\boldsymbol{\beta})\mathscr{T} &= (k\boldsymbol{\alpha}+l\boldsymbol{\beta})\mathscr{T}_A+(k\boldsymbol{\alpha}+l\boldsymbol{\beta})\mathscr{T}_B \\
&= k(\boldsymbol{\alpha}\mathscr{T}_A)+l(\boldsymbol{\beta}\mathscr{T}_A)+k(\boldsymbol{\alpha}\mathscr{T}_B)+l(\boldsymbol{\beta}\mathscr{T}_B) \\
&= k(\boldsymbol{\alpha}\mathscr{T}_A+\boldsymbol{\alpha}\mathscr{T}_B)+l(\boldsymbol{\beta}\mathscr{T}_A+\boldsymbol{\beta}\mathscr{T}_B) \\
&= k(\boldsymbol{\alpha}\mathscr{T})+l(\boldsymbol{\beta}\mathscr{T}),
\end{aligned}
$$

这说明了 $\mathscr{T}$ 是 $\mathfrak{R}_n$ 的线性变换. 很自然地, 我们把这个线性变换 $\mathscr{T}$,称为 $\mathscr{T}_A$ 与 $\mathscr{T}_B$ 的**和**,表为 $\mathscr{T}=\mathscr{T}_A+\mathscr{T}_B$.

下面再研究 $\mathscr{T}$ 所对应的矩阵. 因 $\mathscr{T}_A$ 与 $\mathscr{T}_B$ 的矩阵分别为

$$
\boldsymbol{A}=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} \text{ 与 } \boldsymbol{B}=\begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & & \vdots \\ b_{n1} & b_{n2} & \cdots & b_{nn} \end{pmatrix},
$$

即由定义,有

$$
\begin{cases} \boldsymbol{\alpha}_i\mathscr{T}_A=a_{i1}\boldsymbol{\alpha}_1+a_{i2}\boldsymbol{\alpha}_2+\cdots+a_{in}\boldsymbol{\alpha}_n, \\ \boldsymbol{\alpha}_i\mathscr{T}_B=b_{i1}\boldsymbol{\alpha}_1+b_{i2}\boldsymbol{\alpha}_2+\cdots+b_{in}\boldsymbol{\alpha}_n, \end{cases}
$$

故

$$
\begin{aligned}
\boldsymbol{\alpha}_i\mathscr{T} &= \boldsymbol{\alpha}_i\mathscr{T}_A+\boldsymbol{\alpha}_i\mathscr{T}_B \\
&= (a_{i1}+b_{i1})\boldsymbol{\alpha}_1+(a_{i2}+b_{i2})\boldsymbol{\alpha}_2+\cdots+(a_{in}+b_{in})\boldsymbol{\alpha}_n,
\end{aligned}
$$

这说明 $\mathscr{T}$ 之矩阵的第 i 行第 k 列处的元素 $a_{ik}+b_{ik}$ 等于 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 之相应元素的和. 很自然地, 这样的矩阵叫做矩阵 $\boldsymbol{A}$ 与矩阵 $\boldsymbol{B}$ 的**和**, 表为 $\boldsymbol{A}+\boldsymbol{B}$, 即

$$
\boldsymbol{A}+\boldsymbol{B}=\begin{pmatrix} a_{11}+b_{11} & a_{12}+b_{12} & \cdots & a_{1n}+b_{1n} \\ a_{21}+b_{21} & a_{22}+b_{22} & \cdots & a_{2n}+b_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1}+b_{n1} & a_{n2}+b_{n2} & \cdots & a_{nn}+b_{nn} \end{pmatrix}.
$$

于是有 $\mathscr{T}=\mathscr{T}_{A+B}=\mathscr{T}_A+\mathscr{T}_B$, 这是说: 两个线性变换的和仍是一个线性变换, 它的矩阵等于原来两个线性变换之矩阵的和.

又，容易验证，关于线性变换的加法，有类似于数的加法的下列性质：

交换律　$\mathscr{T}_A+\mathscr{T}_B=\mathscr{T}_B+\mathscr{T}_A$，

结合律　$(\mathscr{T}_A+\mathscr{T}_B)+\mathscr{T}_C=\mathscr{T}_A+(\mathscr{T}_B+\mathscr{T}_C)$.

因之，关于矩阵之运算，也有下述的性质：

$$\boldsymbol{A}+\boldsymbol{B}=\boldsymbol{B}+\boldsymbol{A} \quad 及 \quad (\boldsymbol{A}+\boldsymbol{B})+\boldsymbol{C}=\boldsymbol{A}+(\boldsymbol{B}+\boldsymbol{C}).$$

当然，有关矩阵运算的上面两个等式，也可以由矩阵加法的定义直接推得.

2) 线性变换与数相乘　设有线性变换 $\mathscr{T}_A$，k 为任意数. 令 $\mathfrak{R}_n$ 的任意向量 $\boldsymbol{\alpha}$ 对应于向量 $k(\boldsymbol{\alpha}\mathscr{T}_A)$ 的变换为 $\mathscr{T}$，即 $\boldsymbol{\alpha}\longrightarrow\boldsymbol{\alpha}\mathscr{T}=k(\boldsymbol{\alpha}\mathscr{T}_A)=(\boldsymbol{\alpha}\mathscr{T}_A)k$. 于是，对任意数 a、b，就有

$$\begin{aligned}(a\boldsymbol{\alpha}+b\boldsymbol{\beta})\mathscr{T}&=k[(a\boldsymbol{\alpha}+b\boldsymbol{\beta})\mathscr{T}_A]=k[a(\boldsymbol{\alpha}\mathscr{T}_A)+b(\boldsymbol{\beta}\mathscr{T}_A)]\\&=(ak)(\boldsymbol{\alpha}\mathscr{T}_A)+(bk)(\boldsymbol{\beta}\mathscr{T}_A)\\&=a[k(\boldsymbol{\alpha}\mathscr{T}_A)]+b[k(\boldsymbol{\beta}\mathscr{T}_A)]\\&=a(\boldsymbol{\alpha}\mathscr{T})+b(\boldsymbol{\beta}\mathscr{T}),\end{aligned}$$

这说明了 $\mathscr{T}$ 为 $\mathfrak{R}_n$ 之线性变换，叫做线性变换 $\mathscr{T}_A$ 与数 k 的乘积，表为 $\mathscr{T}=k\cdot\mathscr{T}_A=\mathscr{T}_A\cdot k$.

又因

$$\begin{aligned}\boldsymbol{\alpha}_i(\mathscr{T}_A\cdot k)=(\boldsymbol{\alpha}_i\mathscr{T}_A)k&=(a_{i1}\boldsymbol{\alpha}_1+a_{i2}\boldsymbol{\alpha}_2+\cdots+a_{in}\boldsymbol{\alpha}_n)k\\&=(ka_{i1})\boldsymbol{\alpha}_1+(ka_{i2})\boldsymbol{\alpha}_2+\cdots+(ka_{in})\boldsymbol{\alpha}_n,\end{aligned}$$

故 $\mathscr{T}_A\cdot k=k\cdot\mathscr{T}_A$ 之矩阵中每一元，等于 $\boldsymbol{A}$ 中相应元与数 k 的积，这样的矩阵叫做矩阵 $\boldsymbol{A}$ 与数 k 之积，表以 $k\boldsymbol{A}=\boldsymbol{A}k$，即

$$\boldsymbol{A}k=k\boldsymbol{A}=\begin{pmatrix}ka_{11} & ka_{12} & \cdots & ka_{1n}\\ka_{21} & ka_{22} & \cdots & ka_{2n}\\\vdots & \vdots & & \vdots\\ka_{n1} & ka_{n2} & \cdots & ka_{nn}\end{pmatrix}.$$

因之，线性变换与数之乘积仍是线性变换，其矩阵等于原线性

变换之矩阵与数的乘积.

关于线性变换与数相乘之运算，有如下诸等式:

$$k_1(k_2\mathscr{T}_A)=(k_1k_2)\mathscr{T}_A,\quad 1\cdot\mathscr{T}_A=\mathscr{T}_A,$$

$$(k_1+k_2)\mathscr{T}_A=k_1\mathscr{T}_A+k_2\mathscr{T}_A,$$

$$k(\mathscr{T}_A+\mathscr{T}_B)=k\mathscr{T}_A+k\mathscr{T}_B.$$

因而，也有关于矩阵间相应的诸等式:

$$k_1(k_2\boldsymbol{A})=(k_1k_2)\boldsymbol{A},\quad 1\cdot\boldsymbol{A}=\boldsymbol{A},$$

$$(k_1+k_2)\boldsymbol{A}=k_1\boldsymbol{A}+k_2\boldsymbol{A},\quad k(\boldsymbol{A}+\boldsymbol{B})=k\boldsymbol{A}+k\boldsymbol{B}.$$

这些都很容易验证.

3) 线性变换的乘法　设有 $\mathfrak{R}_n$ 的两个线性变换 $\mathscr{T}_A$ 与 $\mathscr{T}_B$，使 $\mathfrak{R}_n$ 之任意向量 $\boldsymbol{\alpha}$ 对应于向量 $(\boldsymbol{\alpha}\mathscr{T}_A)\mathscr{T}_B$ 之变换 $\mathscr{T}$，即 $\boldsymbol{\alpha}\longrightarrow\boldsymbol{\alpha}\mathscr{T}=(\boldsymbol{\alpha}\mathscr{T}_A)\mathscr{T}_B$，为 $\mathfrak{R}_n$ 的线性变换，这是因为

$$\begin{aligned}(a\boldsymbol{\alpha}+b\boldsymbol{\beta})\mathscr{T}&=[(a\boldsymbol{\alpha}+b\boldsymbol{\beta})\mathscr{T}_A]\mathscr{T}_B\\&=[a(\boldsymbol{\alpha}\mathscr{T}_A)+b(\boldsymbol{\beta}\mathscr{T}_A)]\mathscr{T}_B\\&=a[(\boldsymbol{\alpha}\mathscr{T}_A)]\mathscr{T}_B+b[(\boldsymbol{\beta}\mathscr{T}_A)]\mathscr{T}_B\\&=a(\boldsymbol{\alpha}\mathscr{T})+b(\boldsymbol{\beta}\mathscr{T}).\end{aligned}$$

我们把这个线性变换 $\mathscr{T}$ 叫做先施行 $\mathscr{T}_A$ 后施行 $\mathscr{T}_B$ 的乘积，用符号 $\mathscr{T}_A\mathscr{T}_B$ 表示，即 $\mathscr{T}=\mathscr{T}_A\cdot\mathscr{T}_B$.

又因为

$$\begin{aligned}\boldsymbol{\alpha}_i\mathscr{T}&=\boldsymbol{\alpha}_i(\mathscr{T}_A\mathscr{T}_B)=(\boldsymbol{\alpha}_i\mathscr{T}_A)\mathscr{T}_B\\&=\left(\sum_{j=1}^n a_{ij}\boldsymbol{\alpha}_j\right)\mathscr{T}_B=\sum_{j=1}^n a_{ij}(\boldsymbol{\alpha}_j\mathscr{T}_B)\\&=\sum_{j=1}^n a_{ij}\left(\sum_{k=1}^n b_{jk}\boldsymbol{\alpha}_k\right)=\sum_{k=1}^n\left(\sum_{j=1}^n a_{ij}b_{jk}\right)\boldsymbol{\alpha}_k,\end{aligned}$$

同前面的情况一样，我们把两个线性变换 $\mathscr{T}_A$ 与 $\mathscr{T}_B$ 之积 $\mathscr{T}_A\mathscr{T}_B$ 的矩阵，称为矩阵 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 的积，即

$$\boldsymbol{AB}=\begin{pmatrix} c_{11} & c_{12} & \cdots & c_{1n} \\ c_{21} & c_{22} & \cdots & c_{2n} \\ \vdots & \vdots & & \vdots \\ c_{n1} & c_{n2} & \cdots & c_{nn} \end{pmatrix},$$

这里的 $c_{ik}=\sum_{j=1}^{n} a_{ij}b_{jk}$，即第 i 行第 k 列处的元 c_{ij} 等于矩阵 $\boldsymbol{A}$ 的第 i 行各元与矩阵 $\boldsymbol{B}$ 的第 k 列之各相应元乘积的和．于是，两个线性变换之乘积的矩阵，等于两个线性变换之矩阵的乘积．

容易验证下列关于线性变换之乘法的一些等式：

$$k(\mathscr{T}_A\mathscr{T}_B)=(k\mathscr{T}_A)\mathscr{T}_B,$$
$$(\mathscr{T}_A\mathscr{T}_B)\mathscr{T}_C=\mathscr{T}_A(\mathscr{T}_B\mathscr{T}_C),$$
$$(\mathscr{T}_A+\mathscr{T}_B)\mathscr{T}_C=\mathscr{T}_A\mathscr{T}_C+\mathscr{T}_B\mathscr{T}_C,$$
$$\mathscr{T}_A(\mathscr{T}_B+\mathscr{T}_C)=\mathscr{T}_A\mathscr{T}_B+\mathscr{T}_A\mathscr{T}_C.$$

与此对应，矩阵间的运算也有下列的诸等式：

$$k(\boldsymbol{AB})=(k\boldsymbol{A})\boldsymbol{B},\qquad (\boldsymbol{AB})\boldsymbol{C}=\boldsymbol{A}(\boldsymbol{BC}),$$
$$(\boldsymbol{A}+\boldsymbol{B})\boldsymbol{C}=\boldsymbol{AC}+\boldsymbol{BC},\quad \boldsymbol{A}(\boldsymbol{B}+\boldsymbol{C})=\boldsymbol{AB}+\boldsymbol{AC}.$$

当然，矩阵间的诸等式也可从矩阵乘法与加法的定义直接去验证．

注意：矩阵间的乘法一般不满足交换律，即对两矩阵 $\boldsymbol{A}$ 与 $\boldsymbol{B}$，一般有 $\boldsymbol{AB}\neq\boldsymbol{BA}$．这是矩阵的乘法与数之乘法的差别，这只要看一个例子即可明白．例如

$$\boldsymbol{A}=\begin{pmatrix} 1 & 1 \\ 0 & 2 \end{pmatrix},\qquad \boldsymbol{B}=\begin{pmatrix} 3 & -1 \\ 1 & 0 \end{pmatrix},$$

则

$$\boldsymbol{AB}=\begin{pmatrix} 4 & -1 \\ 2 & 0 \end{pmatrix},\quad \boldsymbol{BA}=\begin{pmatrix} 3 & 1 \\ 1 & 1 \end{pmatrix},$$

故 $\boldsymbol{AB}\neq\boldsymbol{BA}$．正因为这样，一矩阵左乘另一矩阵（即从左边

去乘）和右乘另一矩阵（即从右边去乘）是有区别的．从线性变换的意义看，也能够说明这一点：如在二维空间 $\mathfrak{R}_2$ 中，令 $\mathscr{T}$ 表示使 $\mathfrak{R}_2$ 之每向量（向量的起点固定在原点）以反时针方向绕原点 O 在平面 $\mathfrak{R}_2$ 上旋转 θ 角的变换，$\mathscr{L}$ 表示将 $\mathfrak{R}_2$ 的每向量投影在 x–轴上的变换，显然 $\mathscr{T}$，$\mathscr{L}$ 都是 $\mathfrak{R}_2$ 的线性变换（参见第 1 节）．因之，对于 $\mathscr{T}\mathscr{L}$ 这一变换，就有

$$\begin{aligned}\boldsymbol{\alpha}(\mathscr{T}\mathscr{L})=(\boldsymbol{\alpha}\mathscr{T})\mathscr{L}&=|\boldsymbol{\alpha}\mathscr{T}|\cdot\cos(\theta+\varphi)\cdot\boldsymbol{i}\\&=|\boldsymbol{\alpha}|\cdot\cos(\theta+\varphi)\cdot\boldsymbol{i},\end{aligned}$$

其中 $|\boldsymbol{\beta}|$ 表向量 $\boldsymbol{\beta}$ 的长，φ 为向量 $\boldsymbol{\alpha}$ 与 x–轴之正向所成的角，$\boldsymbol{i}$ 是 x-轴正向上的单位向量，即 $\boldsymbol{\alpha}(\mathscr{T}\mathscr{L})$ 是 x–轴方向的向量．而对 $\mathscr{L}\mathscr{T}$ 这一变换，有

$$\begin{aligned}\boldsymbol{\alpha}(\mathscr{L}\mathscr{T})=(\boldsymbol{\alpha}\mathscr{L})\mathscr{T}=(|\boldsymbol{\alpha}|\cos\varphi\cdot\boldsymbol{i})\mathscr{T}&=|\boldsymbol{\alpha}|\cos\varphi\cdot(\boldsymbol{i}\mathscr{T})\\&=|\boldsymbol{\alpha}|\cos\varphi\cdot(\cos\theta\cdot\boldsymbol{i}+\sin\theta\cdot\boldsymbol{j}),\end{aligned}$$

这里的 $\boldsymbol{j}$ 是 y–轴正向上的单位向量．即 $\boldsymbol{\alpha}(\mathscr{L}\mathscr{T})$ 是和 x–轴夹角为 θ 的向量．显然 $\boldsymbol{\alpha}(\mathscr{T}\mathscr{L})\neq\boldsymbol{\alpha}(\mathscr{L}\mathscr{T})$，故 $\mathscr{T}\mathscr{L}\neq\mathscr{L}\mathscr{T}$．

又，根据矩阵乘积的定义，我们看到，矩阵的乘法规则与行列式的乘法规则完全一样，故有：

定理 两个 n 级矩阵的乘积之行列式等于它们的行列式之乘积．因此，有限多个满秩矩阵的乘积仍然是一个满秩矩阵；在有限多个矩阵中，如果有一个是降秩的，那末它们的乘积也是降秩的．

注意：行列式相乘的法则可有四种（参见第 42 页），所依据的是一个行列式与它的转置行列式相等．可是，对矩阵说来，两个矩阵之积只能如上所述去定义，简言之，就是“行乘列”的规则．因为一矩阵一般不等于它的转置矩阵，故两个矩阵相乘，就不能象行列式那样，也来施行“行乘行”，“列乘行”、或“列乘列”的规律．

第 4 节　特殊的线性变换和矩阵

我们知道，$\mathfrak{R}_n$ 中线性变换 $\mathscr{T}$ 所对应的矩阵，一般来讲是随 $\mathfrak{R}_n$ 之基底的选择不同而不同的. 今问：有没有不管基底的选择怎样，所对应之矩阵恒为一定的线性变换吗？这样的线性变换究竟是怎样的变换呢？

1) 使空间 $\mathfrak{R}_n$ 的每向量都变为 $\mathbf{0}$ 的变换 $\mathscr{T}_O$

易证 $\mathscr{T}_O$ 为 $\mathfrak{R}_n$ 之线性变换：这是因为，从

$$\boldsymbol{\alpha}\mathscr{T}_O=\boldsymbol{\beta}\mathscr{T}_O=(\boldsymbol{\alpha}+\boldsymbol{\beta})\mathscr{T}_O=(k\boldsymbol{\alpha})\mathscr{T}_O=\mathbf{0},$$

当然有

$$(\boldsymbol{\alpha}+\boldsymbol{\beta})\mathscr{T}_O=\boldsymbol{\alpha}\mathscr{T}_O+\boldsymbol{\beta}\mathscr{T}_O,\quad (k\boldsymbol{\alpha})\mathscr{T}_O=k(\boldsymbol{\alpha}\mathscr{T}_O).$$

又，不论 $\boldsymbol{\alpha}_1, \cdots, \boldsymbol{\alpha}_n$ 为 $\mathfrak{R}_n$ 之任何一组基底，由于

$$\boldsymbol{\alpha}_i\mathscr{T}_O=\mathbf{0}=\sum_{j=1}^{n}0\cdot\boldsymbol{\alpha}_j\quad(\text{其中 } i=1,2,\cdots,n),$$

即知 $\mathscr{T}_O$ 所对应之矩阵恒为零矩阵 $\boldsymbol{O}$（即矩阵之元全为数 0），与基底之选择无关.

这一特殊的线性变换 $\mathscr{T}_O$ 叫做零变换，即零变换 $\mathscr{T}_O$ 是我们欲求的一个线性变换.

对于零变换 $\mathscr{T}_O$，易证具有这样的性质：对任意线性变换 $\mathscr{T}_A$ 与任意数 k，恒有

$$\mathscr{T}_A+\mathscr{T}_O=\mathscr{T}_O+\mathscr{T}_A=\mathscr{T}_A,\quad \mathscr{T}_A\mathscr{T}_O=\mathscr{T}_O\mathscr{T}_A=\mathscr{T}_O,$$

以及

$$k\mathscr{T}_O=\mathscr{T}_Ok=\mathscr{T}_O;$$

相应地，关于矩阵的运算，也有

$$\boldsymbol{A}+\boldsymbol{O}=\boldsymbol{O}+\boldsymbol{A}=\boldsymbol{A},\ \boldsymbol{A}\cdot\boldsymbol{O}=\boldsymbol{O}\cdot\boldsymbol{A}=\boldsymbol{O},\ k\boldsymbol{O}=\boldsymbol{O}k=\boldsymbol{O}.$$

这些矩阵间的等式也可以直接由矩阵之加法、乘法、以及与数

之相乘的定义而推出.

若简写 $-\boldsymbol{A}=(-1)\cdot\boldsymbol{A}$, 则 $\boldsymbol{A}-\boldsymbol{B}=\boldsymbol{A}+(-1)\boldsymbol{B}$, 即"减法"有意义, 随而 $\boldsymbol{A}-\boldsymbol{A}=\boldsymbol{O}$. 又, 零矩阵与任何矩阵相乘的结果虽仍为零矩阵, 但两个非零矩阵之积也可能是零矩阵, 这也是矩阵乘法与数量乘法间的不同之点. 例如

$$\boldsymbol{A}=\begin{pmatrix}1 & 1\\ -1 & -1\end{pmatrix},\quad \boldsymbol{B}=\begin{pmatrix}-2 & 1\\ 2 & -1\end{pmatrix},$$

根据矩阵乘法的规律, 显有 $\boldsymbol{AB}=\boldsymbol{O}$. 从线性变换看, 也可得到说明, 如设 $\mathscr{T}$ 与 $\mathscr{S}$ 分别是 $\mathfrak{R}_2$ 中投影在 x–轴上和 y–轴上的投影变换, 显然, 对任 $\boldsymbol{\alpha}\in\mathfrak{R}_2$, 有 $\boldsymbol{\alpha}(\mathscr{T}\mathscr{S})=(\boldsymbol{\alpha}\mathscr{T})\mathscr{S}=\mathbf{0}$, 所以 $\mathscr{T}\mathscr{S}=\mathscr{T}_O$ (零变换), 今后就以 $\mathscr{O}$ 来表示零变换, 即 $\mathscr{T}\mathscr{S}=\mathscr{O}$.

因矩阵之乘法不满足交换律, 故即使 $\boldsymbol{AB}=\boldsymbol{O}$, 也可能会有 $\boldsymbol{BA}\neq\boldsymbol{O}$. 如上面的例子, 就是 $\boldsymbol{AB}=\boldsymbol{O}$, 而 $\boldsymbol{BA}\neq\boldsymbol{O}$. 由于矩阵之乘积有这样的特性, 为以后叙述简便, 引进下面的:

定义 对于矩阵 $\boldsymbol{A}$, 如有非零矩阵 $\boldsymbol{B}$ 存在, 使 $\boldsymbol{AB}$ 或 $\boldsymbol{BA}$ 中有一为零矩阵时, 叫 $\boldsymbol{A}$ 为零因子.

显然, 零矩阵为零因子. 又, 当 $\boldsymbol{A}(\neq\boldsymbol{O})$ 为零因子时, 有矩阵 $\boldsymbol{B}\neq\boldsymbol{O}$ 使 $\boldsymbol{AB}=\boldsymbol{O}$ 或 $\boldsymbol{BA}=\boldsymbol{O}$, 二者必居其一(这二者可能同时成立, 也可能不同时成立). $\boldsymbol{A}$ 为零因子的条件是什么呢? 这可由下面的定理 1 来给出.

定理 1 *n 级矩阵 $\boldsymbol{A}$ 为零因子的充要条件是 $\boldsymbol{A}$ 为降秩的. 并且, 存在 $\boldsymbol{P}\neq\boldsymbol{O}$ 及 $\boldsymbol{Q}\neq\boldsymbol{O}$, 使 $\boldsymbol{AP}=\boldsymbol{O}=\boldsymbol{QA}$.*

证明 设

$$\boldsymbol{A}=\begin{pmatrix}a_{11} & \cdots & a_{1n}\\ \vdots & & \vdots\\ a_{n1} & \cdots & a_{nn}\end{pmatrix}.$$

先证必要性. 若 $\boldsymbol{A}$ 为零因子，则有

$$\boldsymbol{X}=\begin{pmatrix}x_{11} & \cdots & x_{1n}\\ \vdots & & \vdots\\ x_{n1} & \cdots & x_{nn}\end{pmatrix}\neq\boldsymbol{O}$$

使 $\boldsymbol{AX}=\boldsymbol{O}$(或 $\boldsymbol{XA}=\boldsymbol{O}$). 因之，由于至少必有一个 $x_{st}\neq 0$, 故从 $\boldsymbol{AX}=\boldsymbol{O}$ 的第 t 列的元，有

$$\begin{cases}a_{11}x_{1t}+\cdots+a_{1s}x_{st}+\cdots+a_{1n}x_{nt}=0,\\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\ a_{n1}x_{1t}+\cdots+a_{ns}x_{st}+\cdots+a_{nn}x_{nt}=0.\end{cases}$$

表明这齐次方程组有非零解$(x_{1t}, \cdots, x_{st}, \cdots, x_{nt})$, 故其系数之行列式等于零，这就证得了 A 是降秩的.

再证充分性. 若 $\boldsymbol{A}$ 是降秩的，则上齐次方程组必有非零解$(x_1, \cdots, x_n)$，于是，当

$$\boldsymbol{P}=\begin{pmatrix}x_1 & x_1 & \cdots & x_1\\ x_2 & x_2 & \cdots & x_1\\ \vdots & \vdots & & \vdots\\ x_n & x_n & \cdots & x_n\end{pmatrix}$$

时，有 $\boldsymbol{P}\neq\boldsymbol{O}$ 而 $\boldsymbol{AP}=\boldsymbol{O}$, 这就证得了 $\boldsymbol{A}$ 为零因子.

同理，从

$$\sum_{i=1}^{n} a_{ik}x_i=0 \quad (k=1, 2, \cdots, n)$$

存在非零解$(y_1, y_2, \cdots, y_n)$而令

$$\boldsymbol{Q}=\begin{pmatrix}y_1 & y_2 & \cdots & y_n\\ y_1 & y_2 & \cdots & y_n\\ \vdots & \vdots & & \vdots\\ y_1 & y_2 & \cdots & y_n\end{pmatrix}$$

时又有 $\boldsymbol{Q}\neq\boldsymbol{O}$ 及 $\boldsymbol{QA}=\boldsymbol{O}$, 证完.

2) 使空间 $\mathfrak{R}_n$ 之每向量都不变的变换 $\mathscr{T}_E$.

易证 $\mathscr{T}_E$ 是 $\mathfrak{R}_n$ 之线性变换：这是因为

$$(\boldsymbol{\alpha}+\boldsymbol{\beta})\mathscr{T}_E=\boldsymbol{\alpha}+\boldsymbol{\beta}=\boldsymbol{\alpha}\mathscr{T}_E+\boldsymbol{\beta}\mathscr{T}_E$$

及

$$(k\boldsymbol{\alpha})\mathscr{T}_E=k\boldsymbol{\alpha}=k(\boldsymbol{\alpha}\mathscr{T}_E).$$

这一特殊线性变换称为*恒等变换*．不论 $\boldsymbol{\alpha}_1, \cdots, \boldsymbol{\alpha}_n$ 为 $\mathfrak{R}_n$ 之任何基底，由于 $\boldsymbol{\alpha}_i\mathscr{T}_E=\boldsymbol{\alpha}_i$，说明了 $\mathscr{T}_E$ 所对应之矩阵恒为

$$\boldsymbol{E}=\begin{pmatrix}1 & & & 0\\ & 1 & & \\ & & \ddots & \\ 0 & & & 1\end{pmatrix},$$

即主对角线上的都是数 1，而其它位置的全是零，用符号 $\boldsymbol{E}$ 表示之，叫做*单位矩阵*．这里，恒等变换 $\mathscr{T}_E$ 对任何基底而言，所对应的矩阵恒为单位矩阵，故 $\mathscr{T}_E$ 也是我们欲求的线性变换．

容易检验：对任何线性变换 $\mathscr{T}_A$，常有 $\mathscr{T}_A\mathscr{T}_E=\mathscr{T}_E\mathscr{T}_A=\mathscr{T}_A$，于是，对任何矩阵 $\boldsymbol{A}$，恒有 $\boldsymbol{EA}=\boldsymbol{AE}=\boldsymbol{A}$（当然，直接应用矩阵的乘法规则，也可推出这一结论）．这说明，在矩阵乘法运算中，矩阵 $\boldsymbol{E}$ 的作用恰如数 1 在数的乘法运算中的作用．正因为这个原因，所以才叫 $\boldsymbol{E}$ 为单位矩阵．

反之，如有一 n 级矩阵 $\boldsymbol{C}$ 与任何 n 级矩阵 $\boldsymbol{X}$ 相乘恒有 $\boldsymbol{CX}=\boldsymbol{X}$（或都恒有 $\boldsymbol{XC}=\boldsymbol{X}$），则取 $\boldsymbol{X}=\boldsymbol{E}$ 后就有 $\boldsymbol{CE}=\boldsymbol{E}$（或 $\boldsymbol{EC}=\boldsymbol{E}$）；但又 $\boldsymbol{CE}=\boldsymbol{C}$（或 $\boldsymbol{EC}=\boldsymbol{C}$），故有 $\boldsymbol{C}=\boldsymbol{E}$．这就证得了：

定理 2 一 n 级矩阵 $\boldsymbol{C}$ 与任何 n 级矩阵 $\boldsymbol{X}$ 的乘积仍为 $\boldsymbol{X}$ 的充要条件是 $\boldsymbol{C}$ 为单位矩阵．

从线性变换说来，这定理的意义是：恒等变换且只有恒等变换，它与任意线性变换 $\mathscr{T}$ 继续施行的结果仍然是 $\mathscr{T}$．

3) 使空间 $\mathfrak{R}_n$ 的每向量 $\boldsymbol{x}$ 变为 $k\boldsymbol{x}$ (k 为一常数) 的变换 $\mathscr{T}_{kE}$.

不难验证 $\mathscr{T}_{kE}$ 为线性变换: 这是因为,

$$(\boldsymbol{\alpha}+\boldsymbol{\beta})\mathscr{T}_{kE}=k(\boldsymbol{\alpha}+\boldsymbol{\beta})=k\boldsymbol{\alpha}+k\boldsymbol{\beta}=\boldsymbol{\alpha}\mathscr{T}_{kE}+\boldsymbol{\beta}\mathscr{T}_{kE},$$

以及

$$(c\boldsymbol{\alpha})\mathscr{T}_{kE}=k(c\boldsymbol{\alpha})=c(k\boldsymbol{\alpha})=c(\boldsymbol{\alpha}\mathscr{T}_{kE}).$$

这一特殊的线性变换称为相似变换[†]. 容易看出, 关于任何基底, 相似变换所对应的矩阵恒为

$$\begin{pmatrix} k & & & 0 \\ & k & & \\ & & \ddots & \\ 0 & & & k \end{pmatrix},$$

即主对角线上的元都为数 k, 而其他的元全是零. 这矩阵显然等于单位矩阵 $\boldsymbol{E}$ 的 k 倍: $k\boldsymbol{E}$, 这种矩阵称为纯量矩阵. 于是相似变换得表为 $k\mathscr{T}_{E}$, 也是我们所欲求的线性变换之一.

又, 容易获知: 相似变换同任何的线性变换相继施行, 没有施行先后顺序的关系. 换言之, 对任何线性变换 $\mathscr{T}_{A}$ 来说, 恒有

$$k\mathscr{T}_{E}\cdot\mathscr{T}_{A}=\mathscr{T}_{A}\cdot k\mathscr{T}_{E}=k\mathscr{T}_{A}.$$

于是特得

$$k\mathscr{T}_{E}\cdot l\mathscr{T}_{E}=kl\mathscr{T}_{E}.$$

相应地, 对任何矩阵 $\boldsymbol{A}$, 恒有

$$k\boldsymbol{E}\cdot\boldsymbol{A}=\boldsymbol{A}\cdot k\boldsymbol{E}=k\boldsymbol{A},$$

特别地, 有

† 如在 $\mathfrak{R}_2$ 中, $\triangle OAB$ 是由三个向量 $\overrightarrow{OA}$, $\overrightarrow{OB}$ 和 $\overrightarrow{AB}$ 组成. 若将 $\mathfrak{R}_2$ 中每个向量都变为 k 倍, 则 $\overrightarrow{OA}$ 与 $\overrightarrow{OB}$ 分别延长 k 倍而变成了 $\overrightarrow{OA'}$ 与 $\overrightarrow{OB'}$, 于是 $\overrightarrow{A'B'}=k\overrightarrow{AB}$. 因此, 这变换使 $\triangle OAB$ 变成它的相似三角形 $\triangle OA'B'$. 这也是把这样的变换称为相似变换的原因.

$$kE\cdot lE=klE.$$

这些等式说明了：在两个纯量矩阵的积以及纯量矩阵与任意矩阵的积中，矩阵 kE 所起的作用和数 k 所起的作用完全一样，我们叫 kE 为纯量矩阵之意义也在此。

当然，这些矩阵间的等式也可以直接由矩阵相乘的规则而推出。

上面第一个等式告诉我们：每个纯量矩阵同任何一个矩阵 A 相乘为可交换的，它的逆定理是否正确呢？即问：与任何矩阵相乘为可交换的矩阵一定是纯量矩阵吗？答案是肯定的。对此可证明如下：若 n 级矩阵 $C=(c_{ij})$ 与任一个 n 级矩阵 X 相乘为可交换的，则当取

$$X=E_1(2)=\begin{pmatrix}2 & & & 0\\ & 1 & & \\ & & \ddots & \\ 0 & & & 1\end{pmatrix}$$

时（即主对角线上元素顺次为 2, 1, …, 1, 而其它元素为 0），则从 $C\cdot E_1(2)=E_1(2)\cdot C$，即得

$$\begin{pmatrix}2c_{11} & c_{12} & \cdots & c_{1n}\\ 2c_{21} & c_{22} & \cdots & c_{2n}\\ \vdots & \vdots & & \vdots\\ 2c_{n1} & c_{n2} & \cdots & c_{nn}\end{pmatrix}=\begin{pmatrix}2c_{11} & 2c_{12} & \cdots & 2c_{1n}\\ c_{21} & c_{22} & \cdots & c_{2n}\\ \vdots & \vdots & & \vdots\\ c_{n1} & c_{n2} & \cdots & c_{nn}\end{pmatrix},$$

故 $c_{1t}=c_{t1}=0\ (t=2,\ \cdots,\ n)$。同样，取 $X=E_i(2)$ 时（$E_i(2)$ 表示用 2 乘单位矩阵 E 的第 i 行（或列）而得到的 n 级矩阵），利用 $C\cdot E_i(2)=E_i(2)\cdot C$ 又可推得 $c_{it}=c_{ti}=0$（当 $t\neq i$ 时）。

总之，这时就必有 $C=\begin{pmatrix}c_{11} & & & 0\\ & c_{22} & & \\ & & \ddots & \\ 0 & & & c_{nn}\end{pmatrix}$ 形。

再取 $\boldsymbol{X}=\boldsymbol{E}_{12}$($\boldsymbol{E}_{12}$ 是由单位矩阵 $\boldsymbol{E}$ 之第一、二两行互换而得到的矩阵)，则利用 $\boldsymbol{C}\cdot\boldsymbol{E}_{12}=\boldsymbol{E}_{12}\cdot\boldsymbol{C}$，即

$$\begin{pmatrix} 0 & c_{11} & 0 & \cdots & 0 \\ c_{22} & 0 & 0 & \cdots & 0 \\ 0 & 0 & c_{33} & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & c_{nn} \end{pmatrix}=\begin{pmatrix} 0 & c_{22} & 0 & \cdots & 0 \\ c_{11} & 0 & 0 & \cdots & 0 \\ 0 & 0 & c_{33} & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & c_{nn} \end{pmatrix},$$

得知 $c_{11}=c_{22}$. 同样，如取 $\boldsymbol{X}=\boldsymbol{E}_{1i}$(将单位矩阵 $\boldsymbol{E}$ 之第 1、i 两行互换而成)，则又可从 $\boldsymbol{C}\cdot\boldsymbol{E}_{1i}=\boldsymbol{E}_{1i}\cdot\boldsymbol{C}$ 得到 $c_{11}=c_{ii}$. 因之 $c_{11}=c_{22}=\cdots=c_{nn}=k$(令)，则 $\boldsymbol{C}=k\boldsymbol{E}$ 为纯量矩阵. 这就证得了：

定理 3 一 n 级矩阵 $\boldsymbol{C}$ 与任何矩阵 $\boldsymbol{X}$ 相乘为可交换的充要条件是这矩阵为纯量矩阵.

定理 3 的几何意义是：相似变换，而且也只有相似变换，它同任何线性变换相继施行，没有先后顺序的关系.

4) 一般问题，即不论选择怎样的基底，对应之矩阵恒为 $\boldsymbol{A}=(a_{ij})$ 的线性变换 $\mathscr{T}$ 是什么？

先设 $\boldsymbol{\alpha}_i(i=1, 2, \cdots, n)$ 为 $\Re_n$ 之一组基底，由假设可知

$$\boldsymbol{\alpha}_i\mathscr{T}=\sum_{j=1}^{n}a_{ij}\boldsymbol{\alpha}_j, \quad (i=1, 2, \cdots, n). \tag{9}$$

再取 $\Re_n$ 的 n 个向量 $\boldsymbol{\beta}_i(i=1, 2, \cdots, n)$，于是

$$\boldsymbol{\beta}_i=\sum_{j=1}^{n}p_{ij}\boldsymbol{\alpha}_j, \tag{10}$$

由第三章第 2 节知 $\boldsymbol{\beta}_i(i=1, 2, \cdots, n)$ 为 $\Re_n$ 之一组基底的充要条件是

$$\boldsymbol{P}=\begin{pmatrix} p_{11} & \cdots & p_{1n} \\ \vdots & & \vdots \\ p_{n1} & \cdots & p_{nn} \end{pmatrix}$$

为满秩的，即 $\boldsymbol{P}$ 之行列式不等于零.

今设 $\boldsymbol{\beta}_i(i=1, \cdots, n)$ 为 $\mathfrak{R}_n$ 之一组基底，则由假设，又有

$$\boldsymbol{\beta}_i\mathscr{T}=\sum_{j=1}^{n}a_{ij}\boldsymbol{\beta}_j. \tag{11}$$

利用(10)式与(9)式，可知

$$\boldsymbol{\beta}_i\mathscr{T}=\sum_{j=1}^{n}p_{ij}(\boldsymbol{\alpha}_j\mathscr{T})=\sum_{j=1}^{n}p_{ij}\sum_{k=1}^{n}a_{jk}\boldsymbol{\alpha}_k$$

$$=\sum_{k=1}^{n}\left(\sum_{j=1}^{n}p_{ij}a_{jk}\right)\boldsymbol{\alpha}_k,$$

再把(10)式代入(11)式中，又得到

$$\boldsymbol{\beta}_i\mathscr{T}=\sum_{j=1}^{n}a_{ij}\sum_{k=1}^{n}p_{jk}\boldsymbol{\alpha}_k=\sum_{k=1}^{n}\left(\sum_{j=1}^{n}a_{ij}p_{jk}\right)\boldsymbol{\alpha}_k.$$

比较上面的两个等式，并由 $\boldsymbol{\alpha}_1, \cdots, \boldsymbol{\alpha}_n$ 之线性无关性，就得到

$$\sum_{j=1}^{n}p_{ij}a_{jk}=\sum_{j=1}^{n}a_{ij}p_{jk}, \quad (i, k=1, 2, \cdots, n),$$

于是，由矩阵乘法的定义，则知这 n^2 个等式表示着

$$\boldsymbol{PA}=\boldsymbol{AP}. \tag{12}$$

面题设 $\mathscr{T}$ 之矩阵 $\boldsymbol{A}$ 与基底之选择无关，故不论对任何基底 $\boldsymbol{\beta}_i(i=1, 2, \cdots, n)$，恒有(12)式，亦即对任何满秩矩阵 $\boldsymbol{P}$ 恒有(12)式，因之仿照证定理 3 之方法，先取 $\boldsymbol{P}=\boldsymbol{E}_i(2)$，然后再取 $\boldsymbol{P}=\boldsymbol{E}_{1i}$(其中 $i=1, \cdots, n$)，可以推得 $\boldsymbol{A}=k\boldsymbol{E}$ 为纯量矩阵，因面 $\mathscr{T}=k\mathscr{T}_E$ 为相似变换. 特别地，当 k 分别为 1 或 0 时，$\mathscr{T}$ 就分别是恒等变换或零变换. 于是，再结合上述的讨论，就证得了：

定理 4 $\mathfrak{R}_n$ 之线性变换 $\mathscr{T}$ 所对应之矩阵与 $\mathfrak{R}_n$ 之基底的选择无关的充要条件是 $\mathscr{T}$ 只能为零变换，恒等变换，或一般的相似变换.

证明定理 4 的主要关键在(12)式，即 $\boldsymbol{A}$ 与任意满秩矩阵

P 相乘是可交换的，由之推出 $A=kE$ 为纯量矩阵，因而定理 3 中的条件尚可减弱些，实际上，有下面的：

推论 一 n 级矩阵与任何 n 级满秩矩阵相乘为可交换的充要条件是这矩阵为纯量矩阵。

练 习 十 六

1. 不在主对角线上的元都是零的 n 级矩阵，叫对角矩阵。今设 D 是一对角矩阵，它的主对角线上的元 $c_1, c_2, \cdots, c_n$ 两两不同，证明：凡与 D 相乘为可交换的矩阵也一定是对角矩阵。

2. 设 $AX=XA$, $BX=XB$. 证明：$(A+B)X=X(A+B)$, $(AB)X=X(AB)$.

3. $(A+B)(A-B)=A^2-B^2$ 吗？ $(A+B)^2=A^2+2AB+B^2$ 吗？并求这两等式成立的充要条件。

4. 试问：能否有两个 n 级矩阵 A 与 B，使 $AB-BA=E$？

5. 设 C_{ij} 是一个仅在第 i 行第 j 列处的元素为数 1 而其它的元素全是零的 n 级矩阵 $(i, j=1, 2, \cdots, n)$. 试证这 n^2 个 n 级矩阵满足下面的关系式：

$$C_{ij}C_{kl}=\begin{cases} C_{il}, & \text{当 } j=k \text{ 时}, \\ O, & \text{当 } j\neq k \text{ 时}. \end{cases}$$

第 5 节 逆变换、逆矩阵

这一节里还是假定空间 $\mathfrak{R}_n$ 之一组基底 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 固定。所谓线性变换的矩阵，都是指对这组基底而言的。

线性变换与矩阵间既有一一对应关系，为简便计，下面讨论逆变换就着重从矩阵入手。

设 $\mathscr{T}_A$ 是 $\mathfrak{R}_n$ 的线性变换：$\alpha \longrightarrow \alpha'=\alpha\mathscr{T}_A$. 如果 $\alpha' \longrightarrow \alpha$ 也是 $\mathfrak{R}_n$ 的线性变换 $\mathscr{T}_B$，就叫 $\mathscr{T}_B$ 为 $\mathscr{T}_A$ 的逆变换。这时，有 $\mathscr{T}_A \cdot \mathscr{T}_B=\mathscr{T}_B \cdot \mathscr{T}_A=\mathscr{T}_E$. 因此对于矩阵，则有 $AB=BA=E$，而叫 B 为 A 之逆矩阵。因 E 为满秩矩阵，故 A 应是满秩的，这说明了：若 $\mathscr{T}_A$ 有逆变换，则 $\mathscr{T}_A$ 是满秩的。

反过来：当 $\mathscr{T}_A$ 为满秩时，$\mathscr{T}_A$ 有逆变换吗？下面从矩阵的角度来给出回答.

今设

$$A=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}$$

为一满秩矩阵，如有

$$X=\begin{pmatrix} x_{11} & x_{12} & \cdots & x_{1n} \\ x_{21} & x_{22} & \cdots & x_{2n} \\ \vdots & \vdots & & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{nn} \end{pmatrix}$$

满足 $AX=E$，则据矩阵之乘法得知：对于任 j，关于以 x_{1j}，x_{2j}，…，x_{nj} 为未知量的，我们有 n 个方程

$$\begin{cases} a_{11}x_{1j}+a_{12}x_{2j}+\cdots+a_{1n}x_{nj}=0, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{j-1,1}x_{1j}+a_{j-1,2}x_{2j}+\cdots+a_{j-1,n}x_{nj}=0, \\ a_{j1}x_{1j}+a_{j2}x_{2j}+\cdots+a_{jn}x_{nj}=1, \\ a_{j+1,1}x_{1j}+a_{j+1,2}x_{2j}+\cdots+a_{j+1,n}x_{nj}=0, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{n1}x_{1j}+a_{n2}x_{2j}+\cdots+a_{nn}x_{nj}=0. \end{cases}$$

由于这方程组的行列式 $|A|\neq 0$，故据克莱姆定理，知

$$x_{ij}=\frac{1}{|A|}\begin{vmatrix} a_{11} & \cdots & a_{1,i-1} & 0 & a_{1,i+1} & \cdots & a_{1n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{j-1,1} & \cdots & a_{j-1,i-1} & 0 & a_{j-1,i+1} & \cdots & a_{j-1,n} \\ a_{j1} & \cdots & a_{j,i-1} & 1 & a_{j,i+1} & \cdots & a_{jn} \\ a_{j+1,1} & \cdots & a_{j+1,i-1} & 0 & a_{j+1,i+1} & \cdots & a_{j+1,n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n1} & \cdots & a_{n,i-1} & 0 & a_{n,i+1} & \cdots & a_{nn} \end{vmatrix}.$$

再将上行列式依它的第 i 列展开，易知它就是行列式 $|\boldsymbol{A}|$ 中 a_{ji} 的代数余子式 A_{ji}，因此

$$x_{ij}=\frac{1}{|\boldsymbol{A}|}\cdot A_{ji}.\quad (i,\ j=1,\ 2,\ \cdots,\ n)$$

这就说明了：当 $\boldsymbol{A}$ 是满秩时，若它有逆矩阵，它就只有唯一个逆矩阵，即是 $\frac{1}{|\boldsymbol{A}|}\cdot\boldsymbol{A}^*$，式中

$$\boldsymbol{A}^*=\begin{pmatrix} A_{11} & A_{21} & \cdots & A_{n1} \\ A_{12} & A_{22} & \cdots & A_{n2} \\ \vdots & \vdots & & \vdots \\ A_{1n} & A_{2n} & \cdots & A_{nn} \end{pmatrix},$$

其中 A_{ij} 为元 a_{ij} 在行列式 $|\boldsymbol{A}|$ 中的代数余子式.

反之，直接据矩阵之乘法的定义，又易验证：

$$\frac{1}{|\boldsymbol{A}|}\boldsymbol{A}^*\cdot\boldsymbol{A}=\boldsymbol{A}\cdot\frac{1}{|\boldsymbol{A}|}\boldsymbol{A}^*=\boldsymbol{E},$$

即 $\frac{1}{|\boldsymbol{A}|}\boldsymbol{A}^*$ 是满秩矩阵 $\boldsymbol{A}$ 的逆矩阵. 同数的逆一样，矩阵 $\boldsymbol{A}$ 的逆矩阵用 $\boldsymbol{A}^{-1}$ 表示. 于是，得：

定理 1 当且仅当 $\boldsymbol{A}$ 为满秩时，它才有而且只有唯一的逆矩阵

$$\boldsymbol{A}^{-1}=\frac{1}{|\boldsymbol{A}|}\boldsymbol{A}^*.$$

对线性变换来讲，那就是：线性变换 $\mathscr{T}_A$ 有逆变换 $\mathscr{T}_A^{-1}$ 的充要条件是 $\mathscr{T}_A$ 为满秩变换.

注意：如果不要求出逆矩阵的形状是什么，而只是问线性变换 $\mathscr{T}_A$ 有逆变换 $\mathscr{T}_A^{-1}$ 的充要条件，那末不从矩阵入手，直接从线性变换的角度着手，问题解决得还要快些. 事实上，从逆变换的定义即可知道 $\mathscr{T}_A$ 有逆变换 $\mathscr{T}_A^{-1}$ 的充要条件是 $\mathfrak{R}_n$ 之每个向量都有原象且原象只一个（这两个条件有其一就

有其二，参看第2节的推论)，这样，其充要条件为 $\mathfrak{R}=\mathfrak{R}_n\mathscr{T}_A=\mathfrak{R}_n$，即 $\boldsymbol{A}$ 是满秩的，因而 $\mathscr{T}_A$ 是满秩变换.

有了逆矩阵的概念，就可以象数那样来规定矩阵的正、负整指数幂的意义：设 $\boldsymbol{A}$ 为 n 级矩阵，我们规定 $\boldsymbol{A}^0=\boldsymbol{E}$；并当 k 是正整数时，定义 $\boldsymbol{A}^k$ 是 k 个 $\boldsymbol{A}$ 的连乘积 $(\boldsymbol{A}^k=\underbrace{\boldsymbol{A}\boldsymbol{A}\cdots\boldsymbol{A}}_{k\text{个}})$；又若 $\boldsymbol{A}$ 为满秩时，$\boldsymbol{A}^{-k}$ 是 k 个 $\boldsymbol{A}^{-1}$ 的连乘积，即

$$\boldsymbol{A}^{-k}=(\boldsymbol{A}^{-1})^k.$$

同数的情形一样，矩阵的幂指数法则也同样成立：

$$\boldsymbol{A}^{-k}=(\boldsymbol{A}^k)^{-1},\quad \boldsymbol{A}^k\boldsymbol{A}^t=\boldsymbol{A}^{k+t},\quad (\boldsymbol{A}^k)^t=\boldsymbol{A}^{kt},$$

这里的 k, t 为整数(正、负或零)，但只有当 $\boldsymbol{A}$ 是满秩时，$\boldsymbol{A}$ 的负指数幂才有意义.

我们知道，矩阵 $\boldsymbol{A}^*$ 的第 i 行第 j 列处的元 A_{ji} 是行列式 $|\boldsymbol{A}|$ 中第 j 行第 i 列处的元 a_{ji} 的代数余子式. 因 $\boldsymbol{A}^*$ 与 $\boldsymbol{A}$ 之逆矩阵 $\boldsymbol{A}^{-1}$ 有着密切关系，所以 $\boldsymbol{A}^*$ 是一个与 $\boldsymbol{A}$ 有关的矩阵，叫 $\boldsymbol{A}^*$ 为 $\boldsymbol{A}$ 之伴随矩阵. 于是，$\boldsymbol{A}$ 之逆矩阵 $\boldsymbol{A}^{-1}$ 等于其伴随矩阵 $\boldsymbol{A}^*$ 除以 $\boldsymbol{A}$ 之行列式 $|\boldsymbol{A}|$.

再谈 $\boldsymbol{A}^*$ 的一个基本性质.

因 $\boldsymbol{A}\boldsymbol{A}^*=|\boldsymbol{A}|\cdot\boldsymbol{E}$，故 $|\boldsymbol{A}|\cdot|\boldsymbol{A}^*|=|\boldsymbol{A}|^n$. 于是，若 $|\boldsymbol{A}|\neq0$，则有 $|\boldsymbol{A}^*|=|\boldsymbol{A}|^{n-1}$，即 $\boldsymbol{A}$ 为满秩时，$\boldsymbol{A}^*$ 也是满秩的. 但不论 $\boldsymbol{A}$ 是满秩或降秩，伴随矩阵 $\boldsymbol{A}^*$ 是存在的，于是要问：若 $\boldsymbol{A}$ 为降秩时，$\boldsymbol{A}^*$ 怎样呢？显然，当 $\boldsymbol{A}$ 之秩小于 $n-1$ 时，A_{ij} 全是零，因而 $\boldsymbol{A}^*=\boldsymbol{O}$(零矩阵)，当然是降秩的. 然当 $\boldsymbol{A}$ 之秩 $=n-1$ 时，如将 $|\boldsymbol{A}|$ 依第 j 行展开，则从 $|\boldsymbol{A}|=0$ 得知

$$a_{j1}A_{j1}+a_{j2}A_{j2}+\cdots+a_{jn}A_{jn}=0\quad (j=1,\ \cdots,\ n);$$

然由行列式之性质，又有

$$a_{i1}A_{j1}+a_{i2}A_{j2}+\cdots+a_{in}A_{jn}=0\,(i\neq j).$$

这就说明了方程组

$$\begin{cases} a_{11}x_1+a_{12}x_2+\cdots+a_{1n}x_n=0, \\ a_{21}x_1+a_{22}x_2+\cdots+a_{2n}x_n=0, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{n1}x_1+a_{n2}x_2+\cdots+a_{nn}x_n=0 \end{cases} \tag{13}$$

有 $\boldsymbol{A}^*$ 之 n 个列向量为它的解. 但方程组(13)之矩阵 $\boldsymbol{A}$ 之秩 $=n-1$, 故它的任两个解必为线性相关的, 因而 $\boldsymbol{A}^*$ 之任两个列线性相关, 于是其秩不大于1; 然 $\boldsymbol{A}$ 之秩 $=n-1$ 又说明 $\boldsymbol{A}^*$ 中至少有一元素不为零, 故 $\boldsymbol{A}^*$ 之秩又不能小于1. 由是, $\boldsymbol{A}^*$ 之秩等于1, 当然说明了 $\boldsymbol{A}^*$ 是降秩的.

总括上述, 即证得了:

定理 2 设 n 级矩阵 $\boldsymbol{A}$ 与它的伴随矩阵 $\boldsymbol{A}^*$ 的秩分别为 r 与 r^*, 则有:

i) $r=n$ 时, $r^*=n$;

ii) $r=n-1$ 时, $r^*=1$;

iii) $r<n-1$ 时, $r^*=0$.

推论 $|\boldsymbol{A}^*|=|\boldsymbol{A}|^{n-1}$.

证明 $\boldsymbol{A}$ 为满秩时, 这结论在上面已证明了; 当 $\boldsymbol{A}$ 为降秩时, $\boldsymbol{A}^*$ 也因之是降秩的, 故这时 $|\boldsymbol{A}^*|$ 与 $|\boldsymbol{A}|$ 都为零, 所以 $|\boldsymbol{A}^*|=|\boldsymbol{A}|^{n-1}$ 也成立.

由定理1, $\boldsymbol{A}$ 之逆矩阵 $\boldsymbol{A}^{-1}$ 存在的充要条件是 $\boldsymbol{A}$ 为满秩的(即 $|\boldsymbol{A}|\neq 0$), 而这在讲空间 $\mathfrak{R}_n$ 中由一组基底换为另一组基底时就说过, 虽然在那里没有明确地提出逆矩阵这个名词, 在那里是这样说的: 若 $\boldsymbol{\alpha}_i$ $(i=1, 2, \cdots, n)$ 为 $\mathfrak{R}_n$ 之一组基底, 则 $\boldsymbol{\beta}_i$ $(i=1, 2, \cdots, n)$ 亦为 $\mathfrak{R}_n$ 之基底的充要条件是

$$\boldsymbol{\beta}_i=\sum_{j=1}^{n} p_{ij}\boldsymbol{\alpha}_j \quad (i=1, 2, \cdots, n)$$

中的矩阵 $\boldsymbol{P}=(p_{ij})$ 之行列式不等于零，如用这里的术语来表达，那就是 $\boldsymbol{P}$ 之逆矩阵 $\boldsymbol{P}^{-1}$ 存在．于是，自然又会提出下面两个问题：1）当 $\boldsymbol{A}$ 为满秩时，用怎样的方法可判定另一满秩矩阵 $\boldsymbol{B}$ 为 $\boldsymbol{A}$ 的逆矩阵呢？2）给了 n 级满秩矩阵 $\boldsymbol{A}$ 以后，若按照定理 1 去求 $\boldsymbol{A}^{-1}$，就需要计算 n^2 个 $n-1$ 级行列式（伴随矩阵 $\boldsymbol{A}^*$ 的元），计算的工作量太大，于是问求 $\boldsymbol{A}^{-1}$ 的简便方法是什么？

第 1 个问题容易解决：这只要看 $\boldsymbol{AB}$（或 $\boldsymbol{BA}$）$=\boldsymbol{E}$ 能否成立就行了．这是因为，从 $\boldsymbol{AB}=\boldsymbol{E}$（或 $\boldsymbol{BA}=\boldsymbol{E}$）即得 $\boldsymbol{A}^{-1}(\boldsymbol{AB})=\boldsymbol{A}^{-1}\boldsymbol{E}$（或 $(\boldsymbol{BA})\boldsymbol{A}^{-1}=\boldsymbol{EA}^{-1}$），由是据第 4 节定理 2，有 $\boldsymbol{B}=\boldsymbol{BB}=(\boldsymbol{A}^{-1}\boldsymbol{A})\boldsymbol{B}=\boldsymbol{A}^{-1}(\boldsymbol{AB})=\boldsymbol{A}^{-1}\boldsymbol{E}=\boldsymbol{A}^{-1}$（同理，由 $(\boldsymbol{BA})\boldsymbol{A}^{-1}=\boldsymbol{EA}^{-1}$ 也得 $\boldsymbol{B}=\boldsymbol{A}^{-1}$）．故 $\boldsymbol{B}=\boldsymbol{A}^{-1}$ 之充要条件是 $\boldsymbol{AB}=\boldsymbol{E}$（或 $\boldsymbol{BA}=\boldsymbol{E}$）．

从这个判定标准，马上得知下面的

定理 3　(i) 当 $\boldsymbol{A}$ 是满秩时，$(\boldsymbol{A}^{-1})^{-1}=\boldsymbol{A}$；(ii) 当 $\boldsymbol{A}_i(i=1,\cdots,n)$ 都是满秩时，$(\boldsymbol{A}_1\boldsymbol{A}_2\cdots\boldsymbol{A}_k)^{-1}=\boldsymbol{A}_k^{-1}\cdots\boldsymbol{A}_2^{-1}\boldsymbol{A}_1^{-1}$．

证明　$(\boldsymbol{A}^{-1})\boldsymbol{A}=\boldsymbol{E}$ 足以说明 $\boldsymbol{A}=(\boldsymbol{A}^{-1})^{-1}$，即证得了 (i)．又由

$$\begin{aligned}(\boldsymbol{A}_1\boldsymbol{A}_2)(\boldsymbol{A}_2^{-1}\boldsymbol{A}_1^{-1})&=[(\boldsymbol{A}_1\boldsymbol{A}_2)\boldsymbol{A}_2^{-1}]\boldsymbol{A}_1^{-1}\\&=[\boldsymbol{A}_1(\boldsymbol{A}_2\boldsymbol{A}_2^{-1})]\boldsymbol{A}_1^{-1}=(\boldsymbol{A}_1\boldsymbol{E})\boldsymbol{A}_1^{-1}\\&=\boldsymbol{A}_1\boldsymbol{A}_1^{-1}=\boldsymbol{E},\end{aligned}$$

即知 $(\boldsymbol{A}_1\boldsymbol{A}_2)^{-1}=\boldsymbol{A}_2^{-1}\boldsymbol{A}_1^{-1}$，证明了 $k=2$ 时 (ii) 是成立的；再归纳地假定 $(\boldsymbol{A}_1\boldsymbol{A}_2\cdots\boldsymbol{A}_{k-1})^{-1}=\boldsymbol{A}_{k-1}^{-1}\cdots\boldsymbol{A}_2^{-1}\boldsymbol{A}_1^{-1}$，则

$$\begin{aligned}(\boldsymbol{A}_1\boldsymbol{A}_2\cdots\boldsymbol{A}_{k-1}\boldsymbol{A}_k)^{-1}&=[(\boldsymbol{A}_1\boldsymbol{A}_2\cdots\boldsymbol{A}_{k-1})\boldsymbol{A}_k]^{-1}\\&=\boldsymbol{A}_k^{-1}(\boldsymbol{A}_1\boldsymbol{A}_2\cdots\boldsymbol{A}_{k-1})^{-1}\\&=\boldsymbol{A}_k^{-1}\boldsymbol{A}_{k-1}^{-1}\cdots\boldsymbol{A}_2^{-1}\boldsymbol{A}_1^{-1},\end{aligned}$$

即证得了 (ii)．

再解决第2个问题．注意求 A 之逆 A^{-1} 是与单位矩阵 E 紧密地连系的，这是由于 $AA^{-1}=E$ 的原故．又，当 A 满秩时才有逆 A^{-1}，故又牵涉到秩的问题．于是，为了求 A^{-1}，首先要解决 n 级矩阵 A 的秩是 n；而求秩之简便方法是初等变换的方法(参看第二章第4节)，由于这连系着单位矩阵 E，于是就先谈谈单位矩阵 E 经过各种类型的初等变换所产生的效果．

为今后叙述的简洁及引用的方便，先引进：

定义 将单位矩阵 E 施行第1° 类初等变换(例如第 i，j 两行互换)，这样所得的矩阵表为 E_{ij}；对 E 施行第2° 类初等变换(例如第 i 行乘以数 $\mu\neq 0$)，这样所得的矩阵表为 $E_i(\mu)$；对 E 施行第3° 类初等变换(例如第 i 行乘以数 t 后再加到第 j 行)，这样所得的矩阵表为 $E_{ij}(t)$．

注意：符号 E_{ij} 与 $E_i(\mu)$ 在第4节中已有意识地用过．

定义中介绍的这些 E_{ij}、$E_i(\mu)$、$E_{ij}(t)$，统统都叫做初等矩阵．又，由上定义，不难验证：

定理4 (i) E 之第 i，j 两列互换所得的矩阵也是 E_{ij}，并有 $E'_{ij}=E_{ij}=E_{ij}^{-1}$，以及行列式 $|E_{ij}|=-1$．

(ii) E 之第 i 列乘以数 $\mu\neq 0$ 所得的矩阵也是 $E_i(\mu)$，并有 $E_i(\mu)'=E_i(\mu)$，$E_i(\mu)^{-1}=E_i(\mu^{-1})$，及行列式 $|E_i(\mu)|=\mu\neq 0$．

(iii) E 之第 i 列乘以数 t 后再加到第 j 列所得之矩阵却为 $E_{ji}(t)$，即等于 E 之第 j 行乘数 t 后再加到第 i 行上，并有 $E_{ij}(t)'=E_{ji}(t)$，$E_{ij}(t)^{-1}=E_{ij}(-t)$，以及行列式 $|E_{ij}(t)|=1$．

又，据矩阵相乘之定义，可以直接验证：

定理5 (i) 矩阵 A 之第 i，j 两行互换所成之矩阵等于

$E_{ij}A$，A之第i、j两列互换所成之矩阵等于AE_{ij}.

(ii) A之第i行或第i列乘以数μ所成之矩阵分别等于$E_i(\mu)\cdot A$或$A\cdot E_i(\mu)$.

(iii) A之第i行乘以数t后加到第j行所得的矩阵等于$E_{ij}(t)\cdot A$，但A之第i列乘以数t后加到第j列所得的矩阵却等于$A\cdot E_{ji}(t)$.

定理6 任意满秩矩阵都等于有限多个初等矩阵的乘积.

证明 设A是满秩的. 于是，据第二章第4节所说矩阵之标准形得知，可用有限多个初等变换使A变为E，即A与E等价，即$A\simeq E$. 于是，据定理5得知，存在有限多个初等矩阵之积P与有限多个初等矩阵之积Q，使得$PAQ=E$. 但初等矩阵是满秩的，故P与Q亦都为满秩，因而由上式得$A=P^{-1}Q^{-1}$. 然初等矩阵之逆亦为初等矩阵，故P^{-1}与Q^{-1}都等于有限多个初等矩阵之积，因之A也是这样的. 证完.

推论 A与B等价($A\simeq B$)之充要条件是有两个满秩矩阵P与Q，使得$B=PAQ$.

证明 先证必要性. 若$A\simeq B$，则同证定理6一样(即证$A\simeq E$的方法)，存在两个满秩矩阵P与Q，使$PAQ=B$；这就证明了必要条件. 再证充分性. 若存在满秩矩阵P与Q使$B=PAQ$，则因P与Q都等于有限多个初等矩阵之积，而每个初等矩阵左乘或右乘某矩阵就无异乎对该矩阵施行一次"行"的或"列"的初等变换(定理5)，故$B=PAQ$表示着对A施行有限多次的初等变换后而得到B，即$A\simeq B$.

现在来寻求计算逆矩阵的简便方法. 设A是满秩的，于是据第二章第4节所述，只要对矩阵的"行"施行若干次初等变换就能使A变成E. 大致的步骤是：由A之满秩性，知

它的第一列中元素不全为零，故仅施行“行”的初等变换，例如互换某两行，便可使左上角元素不为零；再用适当非零的数乘新矩阵的第一行使左上角等于数 1；然后，用第 i 行减去第一行的若干倍又得到一个与 $\boldsymbol{A}$ 等价的矩阵，其第一列变成了

$$\boldsymbol{e}_1=\begin{pmatrix}1\\0\\\vdots\\0\end{pmatrix}.$$

由于这个与 $\boldsymbol{A}$ 等价的新矩阵亦为满秩的，但它的第一个列向量为 $\boldsymbol{e}_1$，故其第二列中除头一个数外必有非零的数存在，因之再掉换行的顺序可使第二个数不等于 0，然后又可使它等于 1，并再施行若干次第 3° 类型的“行”初等变换，使第二列为

$$\boldsymbol{e}_2=\begin{pmatrix}0\\1\\0\\\vdots\\0\end{pmatrix},$$

而第一列仍为 $\boldsymbol{e}_1$，没有变动．继续这样做下去，仅需利用有限多次的“行”初等变换可使 $\boldsymbol{A}$ 变成 $\boldsymbol{E}$．因之，据定理 5，存在有限多个初等矩阵(如 k 个)$\boldsymbol{P}_1$，$\boldsymbol{P}_2$，…，$\boldsymbol{P}_k$，使得

$$\boldsymbol{P}_1\boldsymbol{P}_2\cdots\boldsymbol{P}_k\boldsymbol{A}=\boldsymbol{E},$$

于是 $\boldsymbol{P}_1\boldsymbol{P}_2\cdots\boldsymbol{P}_k=\boldsymbol{P}_1\boldsymbol{P}_2\cdots\boldsymbol{P}_k\boldsymbol{E}=\boldsymbol{A}^{-1}$，同时等式 $\boldsymbol{P}_1\boldsymbol{P}_2\cdots\boldsymbol{P}_k=\boldsymbol{P}_1\boldsymbol{P}_2\cdots\boldsymbol{P}_k\boldsymbol{E}$ 还说明了这样一个问题，即使 $\boldsymbol{A}$ 变成单位矩阵 $\boldsymbol{E}$ 的一切“行”的初等变换同时也使 $\boldsymbol{E}$ 变成 $\boldsymbol{A}$ 之逆矩阵．这就提供了求逆矩阵的简便方法：先将 $\boldsymbol{E}$ 写在 $\boldsymbol{A}$ 之右边而得一个 $n\times 2n$-矩阵，再对它的行施行若干次初等变换，而使其前 n 个列组成的矩阵 $\boldsymbol{A}$ 变为单位矩阵 $\boldsymbol{E}$，随之后 n 个列所成

之单位矩阵 $\boldsymbol{E}$ 就变成了 $\boldsymbol{A}^{-1}$.

［例］ 求 $\boldsymbol{A}=\begin{pmatrix}1 & 0 & 1\\ 2 & 2 & 0\\ -1 & -\frac{5}{2} & 1\end{pmatrix}$ 之逆 $\boldsymbol{A}^{-1}$.

解 先把 $\boldsymbol{E}$ 写在 $\boldsymbol{A}$ 的右边，使成为一个 3×6–矩阵，再施行"行"的初等变换，就得到

$$\begin{pmatrix}1 & 0 & 1 & 1 & 0 & 0\\ 2 & 2 & 0 & 0 & 1 & 0\\ -1 & -\frac{5}{2} & 1 & 0 & 0 & 1\end{pmatrix}\xrightarrow[\text{第 3 行}+1\times\text{第 1 行}]{\text{第 2 行}-2\times\text{第 1 行}}$$

$$\begin{pmatrix}1 & 0 & 1 & 1 & 0 & 0\\ 0 & 2 & -2 & -2 & 1 & 0\\ 0 & -\frac{5}{2} & 2 & 1 & 0 & 1\end{pmatrix}\xrightarrow{\frac{1}{2}\times\text{第 2 行}}$$

$$\begin{pmatrix}1 & 0 & 1 & 1 & 0 & 0\\ 0 & 1 & -1 & -1 & \frac{1}{2} & 0\\ 0 & -\frac{5}{2} & 2 & 1 & 0 & 1\end{pmatrix}\xrightarrow{\text{第 3 行}+\frac{5}{2}\times\text{第 2 行}}$$

$$\begin{pmatrix}1 & 0 & 1 & 1 & 0 & 0\\ 0 & 1 & -1 & -1 & \frac{1}{2} & 0\\ 0 & 0 & -\frac{1}{2} & -\frac{3}{2} & \frac{5}{4} & 1\end{pmatrix}\xrightarrow{-2\times\text{第 3 行}}$$

$$\begin{pmatrix}1 & 0 & 1 & 1 & 0 & 0\\ 0 & 1 & -1 & -1 & \frac{1}{2} & 0\\ 0 & 0 & 1 & 3 & -\frac{5}{2} & -2\end{pmatrix}\xrightarrow[\text{第 2 行}+1\times\text{第 3 行}]{\text{第 1 行}+(-1)\times\text{第 3 行}}$$

$$\begin{pmatrix} 1 & 0 & 0 & -2 & \frac{5}{2} & 2 \\ 0 & 1 & 0 & 2 & -2 & -2 \\ 0 & 0 & 1 & 3 & -\frac{5}{2} & -2 \end{pmatrix}.$$

于是

$$A^{-1}=\begin{pmatrix} -2 & \frac{5}{2} & 2 \\ 2 & -2 & -2 \\ 3 & -\frac{5}{2} & -2 \end{pmatrix}.$$

同理，将满秩矩阵 A 只施行"列"的初等变换也可使 A 变成 E，同时将单位矩阵 E 施行同样的"列"初等变换后就变成了 A^{-1}. 这时，求 A^{-1} 之方法是把 E 写在 A 的下方而成 6×3-矩阵(上例中 A)，然后再进行"列"初等变换. 对此，读者可作为练习题去做.

练 习 十 七

1. 若 $AX=XA$，A 是满秩，则 $A^{-1}X=XA^{-1}$.

2. 设 $A^2-A+E=O$(零矩阵)，问 A 是满秩还是降秩?

3. 证明：$(-A)^{-1}=-A^{-1}$，$(-A)^*=(-1)^{n-1}A^*$，$(A^*)^*=|A|^{n-2}\cdot A$，其中 n 为矩阵 A 之级.

第 6 节　线性变换在不同基底下所对应矩阵之关系

前面曾指出，线性变换与矩阵间有一一对应关系，这是指对于一组固定的基底 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 而言的. 同一个线性变换，关于不同的基底，它的矩阵可能是互异的，这些矩阵之间有怎样的一种联系呢?

设 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_n$ 与 $\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \cdots, \boldsymbol{\beta}_n$ 是 $\mathfrak{R}_n$ 之两组不同的基底. 线性变换 $\mathscr{T}$ 关于基底 $\boldsymbol{\alpha}_1, \cdots, \boldsymbol{\alpha}_n$ 的矩阵为 $\boldsymbol{A}_{\mathscr{T}}=(a_{ij})$, 关于基底 $\boldsymbol{\beta}_1, \cdots, \boldsymbol{\beta}_n$ 之矩阵为 $\boldsymbol{B}_{\mathscr{T}}=(b_{ij})$, 即

$$\boldsymbol{\alpha}_i\mathscr{T}=a_{i1}\boldsymbol{\alpha}_1+a_{i2}\boldsymbol{\alpha}_2+\cdots+a_{in}\boldsymbol{\alpha}_n,$$
$$\boldsymbol{\beta}_i\mathscr{T}=b_{i1}\boldsymbol{\beta}_1+b_{i2}\boldsymbol{\beta}_2+\cdots+b_{in}\boldsymbol{\beta}_n$$

$(i=1, 2, \cdots, n)$. 因 $\boldsymbol{\alpha}_1, \cdots, \boldsymbol{\alpha}_n$ 是 $\mathfrak{R}_n$ 之基底, 命

$$\boldsymbol{\beta}_i=p_{i1}\boldsymbol{\alpha}_1+p_{i2}\boldsymbol{\alpha}_2+\cdots+p_{in}\boldsymbol{\alpha}_n \quad (i=1, 2, \cdots, n),$$

则

$$\boldsymbol{\beta}_i\mathscr{T}=\left(\sum_{l=1}^{n} p_{il}\boldsymbol{\alpha}_l\right)\mathscr{T}=\sum_{l=1}^{n} p_{il}(\boldsymbol{\alpha}_l\mathscr{T})$$
$$=\sum_{l=1}^{n} p_{il}\sum_{j=1}^{n} a_{lj}\boldsymbol{\alpha}_j=\sum_{j=1}^{n}\left(\sum_{l=1}^{n} p_{il}a_{lj}\right)\boldsymbol{\alpha}_j,$$

又

$$\boldsymbol{\beta}_i\mathscr{T}=\sum_{l=1}^{n} b_{il}\boldsymbol{\beta}_l=\sum_{l=1}^{n} b_{il}\sum_{j=1}^{n} p_{lj}\boldsymbol{\alpha}_j=\sum_{j=1}^{n}\left(\sum_{l=1}^{n} b_{il}p_{lj}\right)\boldsymbol{\alpha}_j,$$

故由于 $\boldsymbol{\alpha}_1, \cdots, \boldsymbol{\alpha}_n$ 线性无关, 就有

$$\sum_{l=1}^{n} p_{il}a_{lj}=\sum_{l=1}^{n} b_{il}p_{lj} \quad (i, j=1, 2, \cdots, n),$$

此即证明了 $\boldsymbol{P}\cdot\boldsymbol{A}_{\mathscr{T}}=\boldsymbol{B}_{\mathscr{T}}\cdot\boldsymbol{P}$, 其中 $\boldsymbol{P}=(p_{ij})$. 因 $\boldsymbol{\beta}_1, \cdots, \boldsymbol{\beta}_n$ 线性无关, 故 $\boldsymbol{P}$ 为满秩, 即 $\boldsymbol{P}^{-1}$ 存在, 由是得

$$\boldsymbol{A}_{\mathscr{T}}=\boldsymbol{P}^{-1}\cdot\boldsymbol{B}_{\mathscr{T}}\cdot\boldsymbol{P}. \tag{14}$$

注意: (14)中的 $\boldsymbol{A}_{\mathscr{T}}$ 与 $\boldsymbol{B}_{\mathscr{T}}$ 是随着线性变换 $\mathscr{T}$ 之变化而变化的, 但矩阵 $\boldsymbol{P}$ 表示由基底 $\boldsymbol{\alpha}_1, \cdots, \boldsymbol{\alpha}_n$ 变为基底 $\boldsymbol{\beta}_1, \cdots, \boldsymbol{\beta}_n$ 的变换的矩阵, 它是不随 $\mathscr{T}$ 而变化的. 适合(14)这种关系的矩阵 $\boldsymbol{A}_{\mathscr{T}}$ 与 $\boldsymbol{B}_{\mathscr{T}}$ 叫做相似矩阵, 用符号 $\boldsymbol{A}_{\mathscr{T}}\sim\boldsymbol{B}_{\mathscr{T}}$ 来表示. 于是有

定理 1 同一个线性变换对于不同的基底的矩阵是相似矩阵.

反之, 相似矩阵也可以认为是同一个线性变换在适当地

选取两组基底后所对应的矩阵，简言之，可以认为表示同一个线性变换.

证明 若对于 $\boldsymbol{A}$，$\boldsymbol{B}$ 有 $\boldsymbol{A}=\boldsymbol{P}^{-1}\boldsymbol{B}\boldsymbol{P}$ 时，则先取任一组基底 $\boldsymbol{\alpha}_i(i=1,2,\cdots,n)$，而作 n 个向量 $\sum_{j=1}^{n}a_{ij}\boldsymbol{\alpha}_j(i=1,2,\cdots,n)$，其中 $(a_{ij})=\boldsymbol{A}$. 于是，据第 2 节的基本定理，可知必有唯一个线性变换 $\mathscr{T}$，使

$$\boldsymbol{\alpha}_i\mathscr{T}=\sum_{j=1}^{n}a_{ij}\boldsymbol{\alpha}_j, \quad (i=1,\ 2,\ \cdots,\ n).$$

再作另 n 个向量 $\boldsymbol{\beta}_i=p_{i1}\boldsymbol{\alpha}_1+\cdots+p_{in}\boldsymbol{\alpha}_n(i=1,\ 2,\ \cdots,\ n)$，其中 $\boldsymbol{P}=(p_{ij})$. 由 $\boldsymbol{P}$ 之满秩性，可知 $\boldsymbol{\beta}_1,\ \cdots,\ \boldsymbol{\beta}_n$ 又为 $\mathfrak{R}_n$ 之一组基底，且有

$$\begin{aligned}\boldsymbol{\beta}_i\mathscr{T}&=\left(\sum_{j=1}^{n}p_{ij}\boldsymbol{\alpha}_j\right)\mathscr{T}=\sum_{j=1}^{n}p_{ij}(\boldsymbol{\alpha}_j\mathscr{T})\\&=\sum_{j=1}^{n}p_{ij}\left(\sum_{k=1}^{n}a_{jk}\boldsymbol{\alpha}_k\right)=\sum_{k=1}^{n}\left(\sum_{j=1}^{n}p_{ij}a_{jk}\right)\boldsymbol{\alpha}_k.\end{aligned}\tag{15}$$

但从 $\boldsymbol{A}=\boldsymbol{P}^{-1}\boldsymbol{B}\boldsymbol{P}$ 得知 $\boldsymbol{P}\boldsymbol{A}=\boldsymbol{B}\boldsymbol{P}$，故

$$\sum_{j=1}^{n}p_{ij}a_{jk}=\sum_{j=1}^{n}b_{ij}p_{jk},$$

式中 $(b_{ij})=\boldsymbol{B}$，将上式代入(15)，又得到

$$\begin{aligned}\boldsymbol{\beta}_i\mathscr{T}&=\sum_{k=1}^{n}\left(\sum_{j=1}^{n}b_{ij}p_{jk}\right)\boldsymbol{\alpha}_k\\&=\sum_{j=1}^{n}b_{ij}\sum_{k=1}^{n}p_{jk}\boldsymbol{\alpha}_k=\sum_{j=1}^{n}b_{ij}\boldsymbol{\beta}_j,\end{aligned}$$

这是说：关于基底 $\boldsymbol{\beta}_i(i=1,\ 2,\ \cdots,\ n)$，线性变换 $\mathscr{T}$ 之矩阵为 $\boldsymbol{B}=(b_{ij})$. 这就证得了 $\mathscr{T}$ 所对应之矩阵可为 $\boldsymbol{A}$ 亦可为 $\boldsymbol{B}$，只是分别关于基底 $\boldsymbol{\alpha}_i(i=1,\ \cdots,\ n)$ 与 $\boldsymbol{\beta}_i(i=1,\ 2,\ \cdots,\ n)$ 而言的. 这就说明：定理 1 之逆成立.

矩阵的相似满足下列三个规律：

1° 反射律：$\boldsymbol{A}\sim\boldsymbol{A}$. (因 $\boldsymbol{A}=\boldsymbol{E}^{-1}\boldsymbol{A}\boldsymbol{E}$，$\boldsymbol{E}$ 为单位矩阵)；

2° 对称律：若 $A \sim B$，则 $B \sim A$.（因由 $A = P^{-1}BP$ 可得 $B = Q^{-1}AQ$，式中 $Q = P^{-1}$）；

3° 传递律：若 $A \sim B$，$B \sim C$，则 $A \sim C$.（因由 $A = P^{-1}BP$，$B = Q^{-1}CQ$，即得 $A = R^{-1}CR$，式中 $R = QP$）.

根据这三条规律，可以把所有的 n 级矩阵加以分类，即将与同一矩阵相似的所有矩阵归于同一个类，因此每个矩阵属于且只属于一类．属于同一个类的矩阵彼此相似，这样的类叫相似类．属于异类的任两个矩阵都不相似．于是相似矩阵所对应的线性变换可以看成相同，唯一的差别只是在基底之选取不同而已．因此，线性变换与矩阵间的一一对应关系，可以进一步表述于下：

定理 2 *n 维空间 $\mathfrak{R}_n$ 中的线性变换与（n 级）矩阵的相似类之间成一一对应关系.*

这里要强调的是：若 $A \sim B$，且 $A \neq B$，对于矩阵 A、B 所对应的线性变换就有两种理解，一种是把它们看成两个不同的线性变换（这是指对于某同一组基底而言，也是我们过去几节研究的问题）；另一种是把它们看成一个相同的线性变换（这是指对于不同的基底而言，亦即这里说的，把一个相似类当作一个整体来看待）．这就反映了应怎样从不同的角度来理解它们的关系.

最后，还要作两点说明.

第一，相似（$A \sim B$）与等价（$A \simeq B$）之间有什么关系？$A \sim B$ 之意义是说有满秩矩阵 P 使 $B = P^{-1}AP$，故由前节定理 6 之推论，可知 $A \simeq B$．反之，$A \simeq B$ 只要求有满秩矩阵 P、Q，使 $B = PAQ$，但并不要求 P 与 Q 互逆，故不一定有 $A \sim B$．这就是说，“相似”一定“等价”，而“等价”却不一定“相似”．既然“相似”必为“等价”，故相似矩阵必有等秩.

注意：相似矩阵之秩相等，也可以从几何的角度来证明．因为当 $\boldsymbol{A}\sim\boldsymbol{B}$ 时，$\boldsymbol{A}$ 与 $\boldsymbol{B}$ 可以看做对于 $\mathfrak{R}_n$ 中适当的两组基底由同一个线性变换 $\mathscr{T}$ 所对应的矩阵，故 $\mathscr{T}$ 变空间 $\mathfrak{R}_n$ 所得的象子空间 $\mathfrak{R}=\mathfrak{R}_n\mathscr{T}$ 之维数一方面等于 $\boldsymbol{A}$ 之秩，另方面又等于 $\boldsymbol{B}$ 之秩．故 $\boldsymbol{A}$ 之秩 $=\boldsymbol{B}$ 之秩．

第二，要说明相似矩阵概念的重要意义．在解析几何里，适当选择坐标系（即变换坐标系）对于几何问题的处理具有很大的作用，例如平移与旋转坐标轴后可使一般的二次方程化为标准形，使我们容易判定这二次方程所表示的究竟是什么曲线或曲面．相应地，在 n 维空间中的问题就是基底的变换．而基底变了后，同一个线性变换（几何问题）对应的矩阵是相似矩阵，为了便于研究线性变换之特征，这就要从相似矩阵类中选取形状最简单的矩阵．所以，所谓矩阵之某某标准形的研究，意思就是要从相似矩阵类中选取形状最简的矩阵．这留在下章里详细讨论．

第 7 节　长方矩阵的运算

前几节研究的线性变换是在同一个向量空间 $\mathfrak{R}_n$ 里的，因而所研究的矩阵都是方阵（n 级的）．但是，研究一个几何问题往往不是在同一个空间内进行的，即要研究从一个空间映射到另一个维数一般是不相同的空间内的几何问题．例如，在图 4-5 中，π 为一给定的平面，l 为一条直线，平面 π 上任一向量 x 投射到直线 l 上的投影变换 $\mathscr{P}$，就是研究两个不同空间 π 与 l 间的几

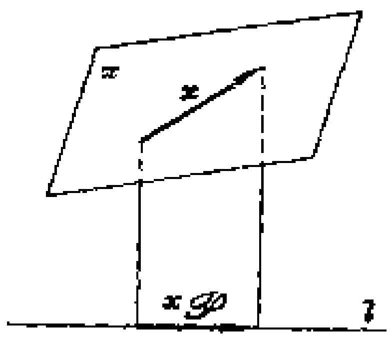


图 4-5

何问题.

今设有一个 m 维空间 $\mathfrak{R}_m$ 及一个 n 维空间 $\mathfrak{R}'_n$，如果有一个确定的方法，能使 $\mathfrak{R}_m$ 的每向量 $\boldsymbol{\alpha}$ 对应于 $\mathfrak{R}'_n$ 的唯一个(由 $\boldsymbol{\alpha}$ 所唯一决定的)向量 $\boldsymbol{\alpha}'$，那末就说 $\mathfrak{R}_m$ 映射在 $\mathfrak{R}'_n$ 内. 用 $\mathscr{T}$ 表示这种映射，就有 $\boldsymbol{\alpha}\mathscr{T}=\boldsymbol{\alpha}'$，即 $\boldsymbol{\alpha}\longrightarrow\boldsymbol{\alpha}'=\boldsymbol{\alpha}\mathscr{T}$，我们也叫 $\mathscr{T}$ 是从 $\mathfrak{R}_m$ 到 $\mathfrak{R}'_n$ 内的变换. 上面的 $\boldsymbol{\alpha}'$ 叫做 $\boldsymbol{\alpha}$ 的象，$\boldsymbol{\alpha}$ 叫做 $\boldsymbol{\alpha}'$ 的原象(关于变换 $\mathscr{T}$ 而言). 有两点值得注意：第一，所谓从 $\mathfrak{R}_m$ 到 $\mathfrak{R}'_n$ 内的变换 $\mathscr{T}$，指的是对于每个 $\boldsymbol{\alpha}\in\mathfrak{R}_m$，都有由 $\boldsymbol{\alpha}$ 确定的唯一个 $\boldsymbol{\alpha}'=\boldsymbol{\alpha}\mathscr{T}\in\mathfrak{R}'_n$；但反之，$\mathfrak{R}'_n$ 之每向量不一定都存在原象(在 $\mathfrak{R}_m$ 中). 第二，据变换 $\mathscr{T}$ 之定义，虽然 $\mathfrak{R}_m$ 的每个向量都有唯一的象在 $\mathfrak{R}'_n$ 内，但 $\mathfrak{R}_m$ 之不同的向量在 $\mathfrak{R}'_n$ 内可能有相同的象，即虽然 $\boldsymbol{\alpha}\neq\boldsymbol{\beta}$，可能有 $\boldsymbol{\alpha}\mathscr{T}=\boldsymbol{\beta}\mathscr{T}$.

如果上述的变换 $\mathscr{T}$ 还满足下列两条件：

i) $(\boldsymbol{\alpha}+\boldsymbol{\beta})\mathscr{T}=\boldsymbol{\alpha}\mathscr{T}+\boldsymbol{\beta}\mathscr{T}$,

ii) $(k\boldsymbol{\alpha})\mathscr{T}=k(\boldsymbol{\alpha}\mathscr{T})$，($k$ 是任意数)

我们就叫 $\mathscr{T}$ 是 $\mathfrak{R}_m$ 到 $\mathfrak{R}'_n$ 内的线性变换. 与第 2 节中讲的一样，得知条件 i 与 ii 是和下面的条件

$$(k\boldsymbol{\alpha}+l\boldsymbol{\beta})\mathscr{T}=k(\boldsymbol{\alpha}\mathscr{T})+l(\boldsymbol{\beta}\mathscr{T})$$

等价的. 例如，图 4-5 中的投影变换 $\mathscr{P}$ 就是 π 到 l 内的线性变换(注意：只要 l 不垂直于 π，投影变换 $\mathscr{P}$ 还是 π 到 l 上的线性变换).

再研究从 $\mathfrak{R}_m$ 到 $\mathfrak{R}'_n$ 内的线性变换与矩阵间的关系. 设 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_m$ 是 $\mathfrak{R}_m$ 的一组固定的基底，$\boldsymbol{\alpha}'_1, \boldsymbol{\alpha}'_2, \cdots, \boldsymbol{\alpha}'_n$ 是 $\mathfrak{R}'_n$ 的一组固定的基底，下面的讨论都是对这两组固定的基底来谈的.

设 $\mathscr{T}$ 是从 $\mathfrak{R}_m$ 到 $\mathfrak{R}'_n$ 内的线性变换. 因 $\boldsymbol{\alpha}_i\mathscr{T}\in\mathfrak{R}'_n$，故由

$$\left\{\begin{array}{l}\boldsymbol{\alpha}_1\mathscr{T}=a_{11}\boldsymbol{\alpha}_1'+a_{12}\boldsymbol{\alpha}_2'+\cdots+a_{1n}\boldsymbol{\alpha}_n',\\ \boldsymbol{\alpha}_2\mathscr{T}=a_{21}\boldsymbol{\alpha}_1'+a_{22}\boldsymbol{\alpha}_2'+\cdots+a_{2n}\boldsymbol{\alpha}_n',\\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\ \boldsymbol{\alpha}_m\mathscr{T}=a_{m1}\boldsymbol{\alpha}_1'+a_{m2}\boldsymbol{\alpha}_2'+\cdots+a_{mn}\boldsymbol{\alpha}_n',\end{array}\right. \tag{16}$$

可得由 $\mathscr{T}$ 所决定的唯一的一个 $m\times n$-矩阵

$$\boldsymbol{A}=\begin{pmatrix}a_{11} & a_{12} & \cdots & a_{1n}\\ a_{21} & a_{22} & \cdots & a_{2n}\\ \vdots & \vdots & & \vdots\\ a_{m1} & a_{m2} & \cdots & a_{mn}\end{pmatrix}.$$

我们叫矩阵 $\boldsymbol{A}$ 为线性变换 $\mathscr{T}$ 所对应之矩阵，也简单地叫做 $\mathscr{T}$ 之矩阵.

反过来，若任给一 $m\times n$-矩阵 $\boldsymbol{A}$（如上述），我们可借(16)式找着一个而且（由下面的定理）也只能找着一个线性变换 $\mathscr{T}$（从 $\mathfrak{R}_m$ 到 $\mathfrak{R}_n'$ 内的），它的矩阵就是 $\boldsymbol{A}$.

定理 1 设 $\boldsymbol{\alpha}_1, \cdots, \boldsymbol{\alpha}_m$ 是 $\mathfrak{R}_m$ 的任一组基底，$\boldsymbol{\alpha}_1', \cdots, \boldsymbol{\alpha}_m'$ 是 $\mathfrak{R}_n'$ 之任 m 个向量，那么，存在一个而且只有一个线性变换（从 $\mathfrak{R}_m$ 到 $\mathfrak{R}_n'$ 内）使 $\boldsymbol{\alpha}_i$ 变为 $\boldsymbol{\alpha}_i'$ $(i=1, \cdots, m)$.

这定理的证法与第 2 节定理的证法完全相同，这里从略.

据定理 1, $\mathfrak{R}_m$ 到 $\mathfrak{R}_n'$ 内的一切线性变换与一切 $m\times n$-矩阵之间形成了一个 1-1 对应关系. 因之，这就可以用矩阵（即一般的 $m\times n$-矩阵）的研究来代替线性变换（从 $\mathfrak{R}_m$ 到 $\mathfrak{R}_n'$ 内）的研究. 应注意的是：这种 1-1 对应关系，是在 $\mathfrak{R}_m$ 与 $\mathfrak{R}_n'$ 之基底都给定的情况来说的. 至于基底变了，它们的关系怎样呢？这就牵涉到长方形矩阵之乘法的问题，这将在下面加以解决.

同第 2 节的讨论一样，易证从 $\mathfrak{R}_m$ 映射到 $\mathfrak{R}_n'$ 内的线性变换 $\mathscr{T}$ 使 $\mathfrak{R}_m$ 中所有向量的象集合 $\mathfrak{R}=\mathfrak{R}_m\mathscr{T}$ 为 $\mathfrak{R}_n'$ 的子空间.

若 $\mathscr{T}$ 之矩阵 $\boldsymbol{A}$ 的秩是 r，显然(16)中 m 个向量 $\boldsymbol{\alpha}_1\mathscr{T}$，…，$\boldsymbol{\alpha}_m\mathscr{T}$ 里只有 r 个是线性无关的，因此 $\mathfrak{N}$ 之维数 $=r$(特别地，当 $\boldsymbol{A}$ 之秩 $=n$ 时，$\mathfrak{N}=\mathfrak{R}_m\mathscr{T}=\mathfrak{R}'_n$，这时当然有 $m\geqslant n$)．又，易证：$\mathfrak{R}_m$ 中之象(经变换 $\mathscr{T}$ 后)等于 $\mathbf{0}(\in\mathfrak{R}'_n)$ 的一切向量的集合 $\mathfrak{M}$ 也是 $\mathfrak{R}_m$ 的子空间，且 $\boldsymbol{x}=\sum\limits_{i=1}^{m}c_i\boldsymbol{\alpha}_i\in\mathfrak{M}$ 之充要条件是

$$\mathfrak{R}'_n\ni\mathbf{0}=\boldsymbol{x}\mathscr{T}=\sum_{i=1}^{m}c_i(\boldsymbol{\alpha}_i\mathscr{T})=\sum_{i=1}^{m}c_i\sum_{j=1}^{n}a_{ij}\boldsymbol{\alpha}'_j$$
$$=\sum_{j=1}^{n}\left(\sum_{i=1}^{m}a_{ij}c_i\right)\boldsymbol{\alpha}'_j,$$

故

$$\sum_{i=1}^{m}a_{ij}c_i=0\quad(j=1,\ 2,\ \cdots,\ n).\tag{17}$$

但因 $\boldsymbol{A}$ 之秩 $=r$，故(17)式有 $m-r$ 个线性无关的解：

$$\boldsymbol{\gamma}_1=(c_1^{(1)},\ \cdots,\ c_m^{(1)}),\ \cdots,\ \boldsymbol{\gamma}_{m-r}=(c_1^{(m-r)},\ \cdots,\ c_m^{(m-r)}),$$

使(17)的任一解 $\boldsymbol{\gamma}=(c_1,\ \cdots,\ c_m)$ 是 $\boldsymbol{\gamma}_1$，…，$\boldsymbol{\gamma}_{m-r}$ 的线性组合，即

$$\boldsymbol{\gamma}=\lambda_1\boldsymbol{\gamma}_1+\cdots+\lambda_{m-r}\boldsymbol{\gamma}_{m-r}.\tag{18}$$

于是，

$$\boldsymbol{x}=\sum_{i=1}^{m}c_i\boldsymbol{\alpha}_i\in\mathfrak{M}$$

的充要条件是存在 $m-r$ 个数 λ_1，λ_2，…，λ_{m-r}，使(18)式成立．然因

$$\boldsymbol{x}_i=\sum_{k=1}^{m}c_k^{(i)}\boldsymbol{\alpha}_k\quad(i=1,\ 2,\ \cdots,\ m-r)$$

为 $\mathfrak{M}$ 中 $m-r$ 个线性无关的向量，故 $\boldsymbol{x}\in\mathfrak{M}$ 的充要条件是 $\boldsymbol{x}$ 表为 $\boldsymbol{x}_1$，$\boldsymbol{x}_2$，…，$\boldsymbol{x}_{m-r}$ 的线性组合：

$$\boldsymbol{x}=\lambda_1\boldsymbol{x}_1+\lambda_2\boldsymbol{x}_2+\cdots+\lambda_{m-r}\boldsymbol{x}_{m-r}.$$

这说明了 $\mathfrak{M}$ 的维数 $=m-r$．总之，$\mathfrak{M}$ 与 $\mathfrak{N}$ 之维数的和等于 m，而 $\mathfrak{N}$ 之维数 $=$ 矩阵 $\boldsymbol{A}$ 之秩．这结果恰是第 2 节定理 1 的

推广，在那里讨论的是在同一个空间内，这里讨论的是在不同空间内．

从 $\mathfrak{R}_m$ 到 $\mathfrak{R}'_n$ 内的线性变换的加法以及与数相乘仍然和第 3 节里面所讨论的完全一样，相应地，矩阵之加法及矩阵与数之相乘，其定义也与以前所说的完全相同，这里就不再叙述了．不过，应注意的是：两个矩阵相加，仅在它们有相同的行数与相同的列数时才有意义，即只有在这样的时候才能定义加法．

再来定义矩阵的乘法．除了上述的两个空间 $\mathfrak{R}_m$ 与 $\mathfrak{R}'_n$ 外，另外还假设有一个 k 维空间 $\mathfrak{R}''_k$，并令 $\boldsymbol{\alpha}''_1, \boldsymbol{\alpha}''_2, \cdots, \boldsymbol{\alpha}''_k$ 是 $\mathfrak{R}''_k$ 的一组基底．当然，还是取 $\mathfrak{R}_m$ 之基底 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_m$ 与 $\mathfrak{R}'_n$ 之基底 $\boldsymbol{\alpha}'_1, \boldsymbol{\alpha}'_2, \cdots, \boldsymbol{\alpha}'_n$ 来讨论．若 $\mathscr{T}$ 是从 $\mathfrak{R}_m$ 到 $\mathfrak{R}'_n$ 内的一个线性变换，它的矩阵为 $m\times n$–矩阵 $\boldsymbol{A}=(a_{ij})$，即有(16)式，再设 $\mathscr{S}$ 是从 $\mathfrak{R}'_n$ 到 $\mathfrak{R}''_k$ 内的一个线性变换，它的矩阵是 $n\times k$–矩阵 $\boldsymbol{B}=(b_{ij})$，即

$$\begin{cases}\boldsymbol{\alpha}'_1\mathscr{S}=b_{11}\boldsymbol{\alpha}''_1+b_{12}\boldsymbol{\alpha}''_2+\cdots+b_{1k}\boldsymbol{\alpha}''_k,\\ \boldsymbol{\alpha}'_2\mathscr{S}=b_{21}\boldsymbol{\alpha}''_1+b_{22}\boldsymbol{\alpha}''_2+\cdots+b_{2k}\boldsymbol{\alpha}''_k,\\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\ \boldsymbol{\alpha}'_n\mathscr{S}=b_{n1}\boldsymbol{\alpha}''_1+b_{n2}\boldsymbol{\alpha}''_2+\cdots+b_{nk}\boldsymbol{\alpha}''_k.\end{cases}\tag{19}$$

如同在第 3 节内讲的一样，易证由 $\boldsymbol{\alpha}(\mathscr{T}\mathscr{S})=(\boldsymbol{\alpha}\mathscr{T})\mathscr{S}$ 所定义的从 $\mathfrak{R}_m$ 到 $\mathfrak{R}''_k$ 内的变换 $\mathscr{T}\mathscr{S}$ 是一个线性变换，叫 $\mathscr{T}\mathscr{S}$ 为 $\mathscr{T}$ 与 $\mathscr{S}$ 的积．又因

$$\begin{aligned}\boldsymbol{\alpha}_i(\mathscr{T}\mathscr{S})&=(\boldsymbol{\alpha}_i\mathscr{T})\mathscr{S}=\left(\sum_{t=1}^{n}a_{it}\boldsymbol{\alpha}'_t\right)\mathscr{S}=\sum_{t=1}^{n}a_{it}(\boldsymbol{\alpha}'_t\mathscr{S})\\&=\sum_{t=1}^{n}a_{it}\left(\sum_{j=1}^{k}b_{tj}\boldsymbol{\alpha}''_j\right)=\sum_{j=1}^{k}\left(\sum_{t=1}^{n}a_{it}b_{tj}\right)\boldsymbol{\alpha}''_j=\sum_{j=1}^{k}p_{ij}\boldsymbol{\alpha}''_j,\end{aligned}$$

这里 $p_{ij}=\sum\limits_{t=1}^{n}a_{it}b_{tj}$，故线性变换 $\mathscr{T}$ 与 $\mathscr{S}$ 之积 $\mathscr{T}\mathscr{S}$ 的矩阵

$\boldsymbol{P}=(p_{ij})$是 $m\times k$-矩阵. 由于 $\boldsymbol{P}$ 对应着线性变换 $\mathscr{T}$ 与 $\mathscr{S}$ 的积，所以也就很自然地叫做线性变换 $\mathscr{T}$ 与 $\mathscr{S}$ 的矩阵 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 的积，用记号 $\boldsymbol{P}=\boldsymbol{AB}$ 表示. 注意：$\boldsymbol{A}$ 有 m 行 n 列，$\boldsymbol{B}$ 有 n 行 k 列，而其积 $\boldsymbol{P}=\boldsymbol{AB}$ 则有 m 行 k 列，它的第 i 行第 j 列处的元素等于 $\boldsymbol{A}$ 之第 i 行的元素与 $\boldsymbol{B}$ 之第 j 列的相应元素的积的和. 所以两个矩阵之积仅在第一个矩阵的列之数目等于第二个矩阵的行之数目时才有意义，即只有这时它们才能相乘.

因此，对于两个一般矩阵 $\boldsymbol{A}$、$\boldsymbol{B}$，即令 $\boldsymbol{AB}$ 有意义，也不一定 $\boldsymbol{BA}$ 有意义. 反之，当 $\boldsymbol{BA}$ 有意义时，不一定 $\boldsymbol{AB}$ 也有意义. 如果 $\boldsymbol{AB}$ 及 $\boldsymbol{BA}$ 都有意义，则 $\boldsymbol{A}$ 之行、列的数目必分别为 $\boldsymbol{B}$ 之列、行的数目.

这里说的矩阵之积的法则和在第 3 节内讲的完全一样，因之两个同级方阵之积的概念是这里的特例. 在第 3 节中关于方阵之积、和、以及与数相乘的那些等式，对于长方阵来说，也都一一成立，这里就不再列出了.

引进了长方矩阵相乘的概念后，对于许多具体问题的处理带来很大的方便，现列举若干例子说明于下：

1）我们曾在第二章里研究过线性方程组

$$\begin{cases} a_{11}x_1+a_{12}x_2+\cdots+a_{1n}x_n=b_1, \\ a_{21}x_1+a_{22}x_2+\cdots+a_{2n}x_n=b_2, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{m1}x_1+a_{m2}x_2+\cdots+a_{mn}x_n=b_m. \end{cases} \tag{20}$$

如果令

$$\boldsymbol{A}=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}, \quad \boldsymbol{X}=\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \quad \boldsymbol{b}=\begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix},$$

则方程组(20)可简写为矩阵的形式:

$$\boldsymbol{A}\boldsymbol{X}=\boldsymbol{b}. \tag{21}$$

于是，解线性方程组(20)的问题，就变成了求(21)中矩阵 $\boldsymbol{X}$ 的问题，这样对于问题的研究当然要简便得多.

特别地，在 $m=n$ 时，$\boldsymbol{A}$ 为 n 级方阵，这时若 $\boldsymbol{A}$ 之行列式 $|\boldsymbol{A}|\neq 0$，则因 $\boldsymbol{A}^{-1}$ 存在，故两边都用 $\boldsymbol{A}^{-1}$ 左乘就得到所求的解为 $\boldsymbol{X}=\boldsymbol{A}^{-1}\boldsymbol{b}$. 这就是说，用解线性方程组的问题就可以通过求逆矩阵来得到解决.

2) 我们也可以将线性变换式(16)改写为矩阵的表示式，如

$$\begin{pmatrix}\boldsymbol{\alpha}_1\mathscr{T}\\ \boldsymbol{\alpha}_2\mathscr{T}\\ \vdots\\ \boldsymbol{\alpha}_m\mathscr{T}\end{pmatrix}=\boldsymbol{A}\begin{pmatrix}\boldsymbol{\alpha}_1'\\ \boldsymbol{\alpha}_2'\\ \vdots\\ \boldsymbol{\alpha}_n'\end{pmatrix}.$$

这样的表写当然简单得多. 但要注意的是：虽然矩阵 $\boldsymbol{A}$ 的元是数，矩阵 $\begin{pmatrix}\boldsymbol{\alpha}_1'\\ \vdots\\ \boldsymbol{\alpha}_n'\end{pmatrix}$ 的元是向量，两者不是同类的东西，可是因数与向量相乘有意义，所以上面的表达式是有意义的，关于这一点，在第1、2节里已有意识地用到过，即将变换公式形式地改写成矩阵的式子.

今引用这样的写法，再度把第6节里面"线性变换在不同基底内所对应矩阵的关系"较简单地导出. 即把第6节中的

$$\begin{aligned}&\boldsymbol{\alpha}_i\mathscr{T}=a_{i1}\boldsymbol{\alpha}_1+a_{i2}\boldsymbol{\alpha}_2+\cdots+a_{in}\boldsymbol{\alpha}_n,\\ &\boldsymbol{\beta}_i\mathscr{T}=b_{i1}\boldsymbol{\alpha}_1+b_{i2}\boldsymbol{\alpha}_2+\cdots+b_{in}\boldsymbol{\alpha}_n,\quad (i=1,\ 2,\ \cdots,\ n)\\ &\boldsymbol{\beta}_i=p_{i1}\boldsymbol{\alpha}_1+p_{i2}\boldsymbol{\alpha}_2+\cdots+p_{in}\boldsymbol{\alpha}_n,\end{aligned}$$

改写为矩阵的等式

$$\begin{pmatrix}\boldsymbol{\alpha}_1\\\boldsymbol{\alpha}_2\\\vdots\\\boldsymbol{\alpha}_n\end{pmatrix}\mathscr{T}=\boldsymbol{A}_{\mathscr{T}}\cdot\begin{pmatrix}\boldsymbol{\alpha}_1\\\boldsymbol{\alpha}_2\\\vdots\\\boldsymbol{\alpha}_n\end{pmatrix},\quad\begin{pmatrix}\boldsymbol{\beta}_1\\\boldsymbol{\beta}_2\\\vdots\\\boldsymbol{\beta}_n\end{pmatrix}\mathscr{T}=\boldsymbol{B}_{\mathscr{T}}\cdot\begin{pmatrix}\boldsymbol{\beta}_1\\\boldsymbol{\beta}_2\\\vdots\\\boldsymbol{\beta}_n\end{pmatrix},$$

$$\begin{pmatrix}\boldsymbol{\beta}_1\\\boldsymbol{\beta}_2\\\vdots\\\boldsymbol{\beta}_n\end{pmatrix}=\boldsymbol{P}\begin{pmatrix}\boldsymbol{\alpha}_1\\\boldsymbol{\alpha}_2\\\vdots\\\boldsymbol{\alpha}_n\end{pmatrix};$$

因 $\boldsymbol{P}$ 为满秩的，由上面的第三个矩阵等式可知

$$\boldsymbol{P}^{-1}\begin{pmatrix}\boldsymbol{\beta}_1\\\boldsymbol{\beta}_2\\\vdots\\\boldsymbol{\beta}_n\end{pmatrix}=\begin{pmatrix}\boldsymbol{\alpha}_1\\\boldsymbol{\alpha}_2\\\vdots\\\boldsymbol{\alpha}_n\end{pmatrix},$$

于是由第一、二个矩阵等式，得

$$\boldsymbol{A}_{\mathscr{T}}\begin{pmatrix}\boldsymbol{\alpha}_1\\\boldsymbol{\alpha}_2\\\vdots\\\boldsymbol{\alpha}_n\end{pmatrix}=\begin{pmatrix}\boldsymbol{\alpha}_1\\\boldsymbol{\alpha}_2\\\vdots\\\boldsymbol{\alpha}_n\end{pmatrix}\mathscr{T}=\boldsymbol{P}^{-1}\begin{pmatrix}\boldsymbol{\beta}_1\\\boldsymbol{\beta}_2\\\vdots\\\boldsymbol{\beta}_n\end{pmatrix}\mathscr{T}$$

$$=\boldsymbol{P}^{-1}\boldsymbol{B}_{\mathscr{T}}\cdot\begin{pmatrix}\boldsymbol{\beta}_1\\\boldsymbol{\beta}_2\\\vdots\\\boldsymbol{\beta}_n\end{pmatrix}=\boldsymbol{P}^{-1}\boldsymbol{B}_{\mathscr{T}}\boldsymbol{P}\begin{pmatrix}\boldsymbol{\alpha}_1\\\boldsymbol{\alpha}_2\\\vdots\\\boldsymbol{\alpha}_n\end{pmatrix},$$

但 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_n$ 线性无关，故

$$\boldsymbol{A}_{\mathscr{T}}=\boldsymbol{P}^{-1}\boldsymbol{B}_{\mathscr{T}}\boldsymbol{P},$$

这就是第 6 节里的公式(14).

又，第 2 节中关于线性变换的两个重要公式，也可以写成矩阵形式．例如向量变换的公式(5)可写为

$$\begin{pmatrix}\boldsymbol{\alpha}_1\\ \boldsymbol{\alpha}_2\\ \vdots\\ \boldsymbol{\alpha}_n\end{pmatrix}\mathscr{T}=\boldsymbol{A}\begin{pmatrix}\boldsymbol{\alpha}_1\\ \boldsymbol{\alpha}_2\\ \vdots\\ \boldsymbol{\alpha}_n\end{pmatrix};$$

表示向量 $\boldsymbol{\alpha}=x_1\boldsymbol{\alpha}_1+\cdots+x_n\boldsymbol{\alpha}_n$，$\boldsymbol{\alpha}\mathscr{T}_A=y_1\boldsymbol{\alpha}_1+\cdots+y_n\boldsymbol{\alpha}_n$ 的坐标变换的公式(6)可写为

$$\begin{pmatrix}y_1\\ y_2\\ \vdots\\ y_n\end{pmatrix}=\boldsymbol{A}'\begin{pmatrix}x_1\\ x_2\\ \vdots\\ x_n\end{pmatrix},$$

或

$$(y_1,\ y_2,\ \cdots,\ y_n)=(x_1,\ x_2,\ \cdots,\ x_n)\boldsymbol{A}.$$

此外，利用矩阵的表示式，还可以表达出同一个向量在不同基底内的坐标变换．例如，设 $\boldsymbol{\gamma}(\in\mathfrak{R}_n)$ 关于基底 $\boldsymbol{\alpha}_1,\ \cdots,\ \boldsymbol{\alpha}_n$ 和 $\boldsymbol{\beta}_1,\ \cdots,\ \boldsymbol{\beta}_n$ 的坐标各为 $x_1,\ \cdots,\ x_n$ 和 $y_1,\ \cdots,\ y_n$，即

$$\begin{aligned}\boldsymbol{\gamma}&=x_1\boldsymbol{\alpha}_1+x_2\boldsymbol{\alpha}_2+\cdots+x_n\boldsymbol{\alpha}_n\\ &=(x_1,\ x_2,\ \cdots,\ x_n)\begin{pmatrix}\boldsymbol{\alpha}_1\\ \boldsymbol{\alpha}_2\\ \vdots\\ \boldsymbol{\alpha}_n\end{pmatrix},\end{aligned}$$

$$\begin{aligned}\boldsymbol{\gamma}&=y_1\boldsymbol{\beta}_1+y_2\boldsymbol{\beta}_2+\cdots+y_n\boldsymbol{\beta}_n\\ &=(y_1,\ y_2,\ \cdots,\ y_n)\begin{pmatrix}\boldsymbol{\beta}_1\\ \boldsymbol{\beta}_2\\ \vdots\\ \boldsymbol{\beta}_n\end{pmatrix},\end{aligned}$$

于是，由

$$\begin{pmatrix}\boldsymbol{\beta}_1\\ \boldsymbol{\beta}_2\\ \vdots\\ \boldsymbol{\beta}_n\end{pmatrix}=\boldsymbol{P}\begin{pmatrix}\boldsymbol{\alpha}_1\\ \boldsymbol{\alpha}_2\\ \vdots\\ \boldsymbol{\alpha}_n\end{pmatrix}$$

可得

$$\boldsymbol{\gamma}=(y_1,\ y_2,\ \cdots,\ y_n)\begin{pmatrix}\boldsymbol{\beta}_1\\ \boldsymbol{\beta}_2\\ \vdots\\ \boldsymbol{\beta}_n\end{pmatrix}=(y_1,\ y_2,\ \cdots,\ y_n)\boldsymbol{P}\begin{pmatrix}\boldsymbol{\alpha}_1\\ \boldsymbol{\alpha}_2\\ \vdots\\ \boldsymbol{\alpha}_n\end{pmatrix},$$

故从 $\boldsymbol{\alpha}_1,\ \boldsymbol{\alpha}_2,\ \cdots,\ \boldsymbol{\alpha}_n$ 之线性无关性，就得

$$(x_1,\ x_2,\ \cdots,\ x_n)=(y_1,\ y_2,\ \cdots,\ y_n)\boldsymbol{P},$$

这就是我们要求的关系，其中 $\boldsymbol{P}$ 为基底变换的矩阵.

3) 长方矩阵相乘的规则可适用于矩阵分块后的乘法. 具体地讲，设 $\boldsymbol{A}=(a_{ij})$ 为 $m\times n$-矩阵，$\boldsymbol{B}=(b_{ij})$ 为 $n\times k$-矩阵，于是 $\boldsymbol{AB}=\boldsymbol{C}=(c_{ij})$ 为 $m\times k$-矩阵. 所谓矩阵分块，是指：

$$\boldsymbol{A}=\begin{pmatrix}\overset{n_1}{\widehat{\boldsymbol{A}_{11}}} & \overset{n_2}{\widehat{\boldsymbol{A}_{12}}} & \cdots & \overset{n_t}{\widehat{\boldsymbol{A}_{1t}}}\\ \boldsymbol{A}_{21} & \boldsymbol{A}_{22} & \cdots & \boldsymbol{A}_{2t}\\ \vdots & \vdots & & \vdots\\ \boldsymbol{A}_{s1} & \boldsymbol{A}_{s2} & \cdots & \boldsymbol{A}_{st}\end{pmatrix}\begin{matrix}\}m_1\\ \}m_2\\ \vdots\\ \}m_s\end{matrix},$$

$$\boldsymbol{B}=\begin{pmatrix}\overset{k_1}{\widehat{\boldsymbol{B}_{11}}} & \overset{k_2}{\widehat{\boldsymbol{B}_{12}}} & \cdots & \overset{k_u}{\widehat{\boldsymbol{B}_{1u}}}\\ \boldsymbol{B}_{21} & \boldsymbol{B}_{22} & \cdots & \boldsymbol{B}_{2u}\\ \vdots & \vdots & & \vdots\\ \boldsymbol{B}_{t1} & \boldsymbol{B}_{t2} & \cdots & \boldsymbol{B}_{tu}\end{pmatrix}\begin{matrix}\}n_1\\ \}n_2\\ \vdots\\ \}n_t\end{matrix},$$

其中 $\boldsymbol{A}_{ij}$ 是 $m_i\times n_j$-矩阵，$\boldsymbol{B}_{ij}$ 是 $n_i\times k_j$-矩阵. $m=m_1+\cdots+m_s$，$n=n_1+\cdots+n_t$，$k=k_1+\cdots+k_u$. 如果记 $m\times k$-矩阵 $\boldsymbol{C}=(c_{ij})$ 为

$$\boldsymbol{C}=\begin{pmatrix}\overset{k_1}{\widehat{\boldsymbol{C}_{11}}} & \overset{k_2}{\widehat{\boldsymbol{C}_{12}}} & \cdots & \overset{k_u}{\widehat{\boldsymbol{C}_{1u}}}\\ \boldsymbol{C}_{21} & \boldsymbol{C}_{22} & \cdots & \boldsymbol{C}_{2u}\\ \vdots & \vdots & & \vdots\\ \boldsymbol{C}_{s1} & \boldsymbol{C}_{s2} & \cdots & \boldsymbol{C}_{su}\end{pmatrix}\begin{matrix}\}m_1\\ \}m_2\\ \vdots\\ \}m_s\end{matrix},$$

其中 $C_{ij}=\sum_{\lambda=1}^{t}A_{i\lambda}B_{\lambda j}$ 为 $m_i\times k_j$-矩阵，这时仍有 $AB=C$. 换句话说，在求矩阵 A 与 B 之积 $AB=C$ 时，可以把 A_{ij} 与 B_{ij} 当做元素看待而按照一般矩阵乘法的规律去进行运算. 对此，留给读者自行验证.

4) 有了长方形矩阵之乘法的意义，就可以解决定理 1 后面所提出的问题：从空间 $\mathfrak{R}_m$ 到空间 $\mathfrak{R}'_n$ 内的线性变换与 $m\times n$-矩阵间的 1-1 对应关系，当基底发生变化时，这 1-1 对应关系又将怎样呢？

设 $\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_m$ 与 $\boldsymbol{\beta}_1,\cdots,\boldsymbol{\beta}_m$ 都是 $\mathfrak{R}_m$ 的基底，于是

$$\begin{pmatrix}\boldsymbol{\beta}_1\\ \vdots\\ \boldsymbol{\beta}_m\end{pmatrix}=P\begin{pmatrix}\boldsymbol{\alpha}_1\\ \vdots\\ \boldsymbol{\alpha}_m\end{pmatrix},\quad P=(p_{ij})$$

是 m 级满秩矩阵. 又若 $\boldsymbol{\alpha}'_1,\cdots,\boldsymbol{\alpha}'_n$ 与 $\boldsymbol{\beta}'_1,\cdots,\boldsymbol{\beta}'_n$ 为 $\mathfrak{R}'_n$ 之两组基底，则

$$\begin{pmatrix}\boldsymbol{\beta}'_1\\ \vdots\\ \boldsymbol{\beta}'_n\end{pmatrix}=Q\begin{pmatrix}\boldsymbol{\alpha}'_1\\ \vdots\\ \boldsymbol{\alpha}'_n\end{pmatrix},$$

而 $Q=(q_{ij})$ 为 n 级满秩矩阵. 并设从 $\mathfrak{R}_m$ 到 $\mathfrak{R}'_n$ 内之线性变换 $\mathscr{T}$ 关于基底 $\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_m$ 与 $\boldsymbol{\alpha}'_1,\cdots,\boldsymbol{\alpha}'_n$ 所对应之 $m\times n$-矩阵为 $A_{\mathscr{T}}$，而关于基底 $\boldsymbol{\beta}_1,\cdots,\boldsymbol{\beta}_m$ 与 $\boldsymbol{\beta}'_1,\cdots,\boldsymbol{\beta}'_n$ 所对应的 $m\times n$-矩阵为 $B_{\mathscr{T}}$，即

$$\begin{pmatrix}\boldsymbol{\alpha}_1\\ \vdots\\ \boldsymbol{\alpha}_m\end{pmatrix}\mathscr{T}=A_{\mathscr{T}}\cdot\begin{pmatrix}\boldsymbol{\alpha}'_1\\ \vdots\\ \boldsymbol{\alpha}'_n\end{pmatrix},\quad \begin{pmatrix}\boldsymbol{\beta}_1\\ \vdots\\ \boldsymbol{\beta}_m\end{pmatrix}\mathscr{T}=B_{\mathscr{T}}\cdot\begin{pmatrix}\boldsymbol{\beta}'_1\\ \vdots\\ \boldsymbol{\beta}'_n\end{pmatrix}.$$

故有

$$PA_{\mathscr{T}}\begin{pmatrix}\boldsymbol{\alpha}'_1\\ \vdots\\ \boldsymbol{\alpha}'_n\end{pmatrix}=P\cdot\begin{pmatrix}\boldsymbol{\alpha}_1\\ \vdots\\ \boldsymbol{\alpha}_m\end{pmatrix}\mathscr{T}=\begin{pmatrix}\boldsymbol{\beta}_1\\ \vdots\\ \boldsymbol{\beta}_m\end{pmatrix}\mathscr{T}$$

$$=\boldsymbol{B}_{\mathscr{T}}\cdot\begin{pmatrix}\boldsymbol{\beta}_1'\\ \vdots\\ \boldsymbol{\beta}_n'\end{pmatrix}=\boldsymbol{B}_{\mathscr{T}}\boldsymbol{Q}\cdot\begin{pmatrix}\boldsymbol{\alpha}_1'\\ \vdots\\ \boldsymbol{\alpha}_n'\end{pmatrix},$$

因之，由于 $\boldsymbol{\alpha}_1', \cdots, \boldsymbol{\alpha}_n'$ 的线性无关性，可知 $\boldsymbol{P}\boldsymbol{A}_{\mathscr{T}}=\boldsymbol{B}_{\mathscr{T}}\boldsymbol{Q}$，即

$$\boldsymbol{A}_{\mathscr{T}}=\boldsymbol{P}^{-1}\boldsymbol{B}_{\mathscr{T}}\boldsymbol{Q}. \tag{22}$$

(22)式即为我们所求之关系．这就是说：从 $\mathfrak{R}_m$ 到 $\mathfrak{R}_n'$ 内的同一个线性变换 $\mathscr{T}$ 关于各个空间的不同基底所对应之 $m\times n$-矩阵 $\boldsymbol{A}_{\mathscr{T}}$ 与 $\boldsymbol{B}_{\mathscr{T}}$ 之关系，是存在与 $\mathscr{T}$ 无关的两个矩阵 $\boldsymbol{P}$ 和 $\boldsymbol{Q}$（分别为 m 级与 n 级的）使(22)式成立．反之，如有两个满秩矩阵 $\boldsymbol{P}$ 与 $\boldsymbol{Q}$ 使(22)成立，则 $\boldsymbol{A}_{\mathscr{T}}$ 与 $\boldsymbol{B}_{\mathscr{T}}$ 必表同一个线性变换．特别地，当 $\mathfrak{R}_n'=\mathfrak{R}_m$（因之 $m=n$），并令 $\boldsymbol{\alpha}_i'=\boldsymbol{\alpha}_i$ 及 $\boldsymbol{\beta}_i'=\boldsymbol{\beta}_i$ 时，(22)式就变成了(14)式．这就是说，第 6 节中所谈的问题即为这里的一个特例．这样，(22)式的几何意义简单说来是表示同一个线性变换，其代数意义是说明两个 $m\times n$-矩阵为等价的．

我们已介绍了矩阵之加法、减法及乘法的运算，于是自然会提出矩阵之除法的问题．当然，除法应为乘法的逆运算，但因矩阵的乘法一般不满足交换律，故定义除法就应有两种：一为左除，即求矩阵 $\boldsymbol{X}$ 使 $\boldsymbol{AX}=\boldsymbol{B}$；另一为右除，即求矩阵 $\boldsymbol{Y}$ 使 $\boldsymbol{YA}=\boldsymbol{B}$．如果 $\boldsymbol{A}$ 为 $m\times k$-矩阵，$\boldsymbol{B}$ 为 $n\times l$-矩阵，那末 $\boldsymbol{X}$ 存在的必要条件是 $m=n$，$\boldsymbol{Y}$ 存在的必要条件是 $k=l$；但即使 $m=n$, $k=l$, $\boldsymbol{X}$ 与 $\boldsymbol{Y}$ 也未必存在，因为这不是充分条件．特别地，当 $\boldsymbol{A}$、$\boldsymbol{B}$ 都是 n 级矩阵时，若 $\boldsymbol{A}$ 是降秩的，$\boldsymbol{B}$ 为满秩，则显然 $\boldsymbol{X}$ 与 $\boldsymbol{Y}$ 都不存在；但若 $\boldsymbol{A}$, $\boldsymbol{B}$ 都是降秩的，那末 $\boldsymbol{X}$、$\boldsymbol{Y}$ 中，或两者存在，或仅一存在，或都不存在，且即使存在，大都也不是唯一的，因此就无法谈 $\boldsymbol{A}$ 除 $\boldsymbol{B}$ 的问题了．

可是，当 $\boldsymbol{A}$ 为满秩 n 级矩阵，则 $\boldsymbol{A}^{-1}$ 存在，故由 $\boldsymbol{AX}=\boldsymbol{B}$

与 $\boldsymbol{YA}=\boldsymbol{B}$ 分别知必有 $\boldsymbol{X}=\boldsymbol{A}^{-1}\boldsymbol{B}$ 与 $\boldsymbol{Y}=\boldsymbol{BA}^{-1}$；反之，$\boldsymbol{X}=\boldsymbol{A}^{-1}\boldsymbol{B}$ 与 $\boldsymbol{Y}=\boldsymbol{BA}^{-1}$ 的确又分别是 $\boldsymbol{AX}=\boldsymbol{B}$ 与 $\boldsymbol{YA}=\boldsymbol{B}$ 之解（这可直接代入检验之）．因此就有下面的

定理 2 设 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 都是 n 级矩阵，且 $\boldsymbol{A}$ 是满秩的，那就有一个且只一个矩阵 $\boldsymbol{X}=\boldsymbol{A}^{-1}\boldsymbol{B}$ 适合 $\boldsymbol{AX}=\boldsymbol{B}$. 同样，$\boldsymbol{YA}=\boldsymbol{B}$ 也有唯一个解 $\boldsymbol{Y}=\boldsymbol{BA}^{-1}$.

再谈转置矩阵的问题．显然，$m\times n$–矩阵的转置矩阵是 $n\times m$–矩阵．对于乘法来说，有

定理 3 若 k 个矩阵 $\boldsymbol{A}_1, \boldsymbol{A}_2, \cdots, \boldsymbol{A}_k$ 的积 $\boldsymbol{A}_1\boldsymbol{A}_2\cdots\boldsymbol{A}_k$ 有意义，则 $(\boldsymbol{A}_1\boldsymbol{A}_2\cdots\boldsymbol{A}_k)'=\boldsymbol{A}_k'\cdots\boldsymbol{A}_2'\boldsymbol{A}_1'$（即矩阵之积的转置等于各个转置之积但排以逆序）．

证明 当 $\boldsymbol{A}_1\boldsymbol{A}_2\cdots\boldsymbol{A}_k$ 有意义时，不难检验 $(\boldsymbol{A}_1\boldsymbol{A}_2\cdots\boldsymbol{A}_k)'$ 与 $\boldsymbol{A}_k'\cdots\boldsymbol{A}_2'\boldsymbol{A}_1'$ 都有意义，且有相同的行数与列数．于是，只需证明它们相等．若令 $\boldsymbol{A}_t=(a_{ij}^{(t)})$，则 $(\boldsymbol{A}_1\boldsymbol{A}_2)'$ 之第 i 行第 j 列处的元 $=\boldsymbol{A}_1\boldsymbol{A}_2$ 之第 j 行第 i 列处的元 $=\sum a_{j\lambda}^{(1)}a_{\lambda i}^{(2)}=\sum a_{\lambda i}^{(2)}a_{j\lambda}^{(1)}$ $=\boldsymbol{A}_2$ 之第 i 列的各元与 $\boldsymbol{A}_1$ 之第 j 行的各相应元之积的和 $=\boldsymbol{A}_2'$ 之第 i 行各元与 $\boldsymbol{A}_1'$ 之第 j 列的各相应元之积的和 $=\boldsymbol{A}_2'\boldsymbol{A}_1'$ 之第 i 行第 j 列处的元，故 $(\boldsymbol{A}_1\boldsymbol{A}_2)'=\boldsymbol{A}_2'\boldsymbol{A}_1'$，即证明了 $k=2$ 时定理 3 是成立的．

于是，再归纳地假定 $(\boldsymbol{A}_1\boldsymbol{A}_2\cdots\boldsymbol{A}_{k-1})'=\boldsymbol{A}_{k-1}'\cdots\boldsymbol{A}_2'\boldsymbol{A}_1'$ 之后，就有

$$\begin{aligned}(\boldsymbol{A}_1\boldsymbol{A}_2\cdots\boldsymbol{A}_{k-1}\boldsymbol{A}_k)'&=[(\boldsymbol{A}_1\boldsymbol{A}_2\cdots\boldsymbol{A}_{k-1})\boldsymbol{A}_k]'\\&=\boldsymbol{A}_k'(\boldsymbol{A}_1\boldsymbol{A}_2\cdots\boldsymbol{A}_{k-1})'=\boldsymbol{A}_k'\boldsymbol{A}_{k-1}'\cdots\boldsymbol{A}_2'\boldsymbol{A}_1'.\end{aligned}$$

这就说明了定理 3 在 $k-1$ 个矩阵的场合若成立，则在 k 个矩阵的情形时也必成立．故由归纳法完全证明了定理 3.

本节的几何问题再作深入一步地探讨于下．

设 $\mathfrak{R}_n$ 为一般的 n 维复空间（对于实空间也同样成立），

而$\mathfrak{C}$为一切复数之集合，于是$\mathfrak{C}$自身也可以视为一个复空间，其维数等于1. 取$\mathfrak{R}_n$之一组基底$\boldsymbol{e}_1, \boldsymbol{e}_2, \cdots, \boldsymbol{e}_n$及$\mathfrak{C}$之基底1. 考虑$\mathfrak{R}_n$到$\mathfrak{C}$内的一切线性变换之集合$\mathfrak{R}^*$. 根据定理1，不论$\alpha_1, \alpha_2, \cdots, \alpha_n$为$\mathfrak{C}$中任$n$个数，一定有一个且只一个线性变换$\mathscr{T}$使$\boldsymbol{e}_i\mathscr{T}=\alpha_i$，且关于$\mathfrak{R}_n$及$\mathfrak{C}$之基底$\boldsymbol{e}_1, \cdots, \boldsymbol{e}_n$及1说来，$\mathscr{T}$所对应之矩阵为

$$\begin{pmatrix}\alpha_1\\ \alpha_2\\ \vdots\\ \alpha_n\end{pmatrix}.$$

再作$\mathfrak{R}_n$到$\mathfrak{C}$内的n个特殊线性变换$\mathscr{T}_1^*, \mathscr{T}_2^*, \cdots, \mathscr{T}_n^*$，使

$$\boldsymbol{e}_i\mathscr{T}_j^*=\delta_{ij}=\begin{cases}1, & \text{当 } i=j \text{ 时;}\\ 0, & \text{当 } i\neq j \text{ 时,}\end{cases}$$

即

$$\begin{cases}\boldsymbol{e}_1\mathscr{T}_1^*=1,\\ \boldsymbol{e}_2\mathscr{T}_1^*=0,\\ \quad\vdots\\ \boldsymbol{e}_n\mathscr{T}_1^*=0;\end{cases}\quad \begin{cases}\boldsymbol{e}_1\mathscr{T}_2^*=0,\\ \boldsymbol{e}_2\mathscr{T}_2^*=1,\\ \quad\vdots\\ \boldsymbol{e}_n\mathscr{T}_2^*=0;\end{cases}\quad\cdots\quad \begin{cases}\boldsymbol{e}_1\mathscr{T}_n^*=0,\\ \quad\vdots\\ \boldsymbol{e}_{n-1}\mathscr{T}_n^*=0,\\ \boldsymbol{e}_n\mathscr{T}_n^*=1.\end{cases}$$

不难验证：若$\boldsymbol{x}=\sum\limits_1^n k_i\boldsymbol{e}_i\in\mathfrak{R}_n$，则

$$\begin{aligned}\boldsymbol{x}\left(\sum_{j=1}^n\alpha_j\mathscr{T}_j^*\right)&=\sum_{i=1}^n k_i\cdot\boldsymbol{e}_i\left(\sum_{j=1}^n\alpha_j\mathscr{T}_j^*\right)\\ &=\sum_{i=1}^n k_i[\boldsymbol{e}_i(\alpha_i\mathscr{T}_i^*)]=\sum_{i=1}^n k_i\alpha_i.\end{aligned}$$

但又因

$$\boldsymbol{x}\mathscr{T}=\left(\sum_{i=1}^n k_i\boldsymbol{e}_i\right)\mathscr{T}=\sum_{i=1}^n k_i(\boldsymbol{e}_i\mathscr{T})=\sum_{i=1}^n k_i\alpha_i,$$

故有

$$\mathscr{T}=\sum_{j=1}^n\alpha_j\mathscr{T}_j^*,$$

即 $\mathfrak{R}_n$ 到 $\mathfrak{C}$ 内之任一线性变换 $\mathscr{T}$ 可写为 $\mathscr{T}_1^*, \mathscr{T}_2^*, \cdots, \mathscr{T}_n^*$ 之线性组合. 且若 $\sum_{j=1}^{n}\lambda_j\mathscr{T}_j^*=\mathscr{O}$ (零变换), 则

$$\mathscr{O}=e_i\left(\sum_{j=1}^{n}\lambda_j\mathscr{T}_j^*\right)=\sum_{j=1}^{n}\lambda_j(e_i\mathscr{T}_j^*)=\lambda_i(e_i\mathscr{T}_i^*)=\lambda_i,$$

这就证明了 $\mathscr{T}_1^*, \cdots, \mathscr{T}_n^*$ 线性无关. 所以, $\mathfrak{R}^*$ 实际上是一个 n 维的向量空间, 而 $\mathscr{T}_1^*, \cdots, \mathscr{T}_n^*$ 为其一组基底. 今后, 记 $\mathfrak{R}^*=\mathfrak{R}_n^*$, 并给出下面的

定义 叫上面的 $\mathfrak{R}^*=\mathfrak{R}_n^*$ 为空间 $\mathfrak{R}_n$ 的共轭空间, 而叫 $(\mathscr{T}_1^*, \mathscr{T}_2^*, \cdots, \mathscr{T}_n^*)$ 为相对于 $\mathfrak{R}_n$ 之基底 $(e_1, e_2, \cdots, e_n)$ 的 $\mathfrak{R}_n^*$ 之补基底.

今令 a 为 $\mathfrak{R}_n$ 之一固定向量, 而让 $\mathscr{T}$ 取遍共轭空间 $\mathfrak{R}_n^*$. 于是 $\mathscr{T}\to a\mathscr{T}$ 为 $\mathfrak{R}_n^*$ 在 $\mathfrak{C}$ 内的一个映射. 由于 $a(\mathscr{T}+\mathscr{S})=a\mathscr{T}+a\mathscr{S}$ 及 $a(k\mathscr{T})=k(a\mathscr{T})$, 所以映射 $\mathscr{T}\to a\mathscr{T}$ 实际上是 $\mathfrak{R}_n^*$ 到 $\mathfrak{C}$ 内的一个线性变换. 也就是说, $\mathfrak{R}_n$ 的每元 a 得决定 $\mathfrak{R}_n$ 之共轭空间 $\mathfrak{R}_n^*$ 的一个线性变换 (在 $\mathfrak{C}$ 内的), 即每 $a(\in\mathfrak{R}_n)$ 作用在 $\mathfrak{R}_n^*$ 上是起到了 $\mathfrak{R}_n^*$ 在 $\mathfrak{C}$ 内的线性变换的功用, 即每 $a(\in\mathfrak{R}_n)$ 可看作是 $\mathfrak{R}_n^*$ 之共轭空间 $\mathfrak{R}_n^{**}$ 的一元. 又, 当 $a\neq b$ 时, 易知 a 与 b 也为 $\mathfrak{R}_n^{**}$ 之不同元(这是因为, $a\mathscr{T}=b\mathscr{T}$ 对一切 $\mathscr{T}\in\mathfrak{R}_n^*$ 成立的充要条件是 $(a-b)\mathscr{T}=0$ 对一切 $\mathscr{T}\in\mathfrak{R}_n^*$ 成立, 故有 $a=b$).

由于"线性变换"的符号总是写在右边, 因此根据上述如将 $\mathfrak{R}_n$ 之元 a 视为 $\mathfrak{R}_n^*$ 在 $\mathfrak{C}$ 内的线性变换(即视 a 为 $\mathfrak{R}_n^{**}$ 之元), 就应习惯地把 a 写在 $\mathfrak{R}_n^*$ 之元的右边, 即写 $a\mathscr{T}=\mathscr{T}a$. 但 $a\mathscr{T}=\mathscr{T}a\in\mathfrak{C}$, 而 $a\mathscr{T}$ 是把 $\mathscr{T}$ 当作 $\mathfrak{R}_n$ 在 $\mathfrak{C}$ 内的线性变换, $\mathscr{T}a$ 是把 a 当作 $\mathfrak{R}_n^*$ 在 $\mathfrak{C}$ 内的线性变换, 也是说: 习惯上把 $a\mathscr{T}$ 之 a 看作是 $\mathfrak{R}_n$ 之元, 把 $\mathscr{T}a$ 之 a 看作是 $\mathfrak{R}_n^{**}$ 之元.

因之，映射 $\boldsymbol{a}\to\mathscr{T}\boldsymbol{a}$ 就可看作是 $\mathfrak{R}_n$ 到 $\mathfrak{R}_n^{**}$ 内的映射（但这里的 $\mathscr{T}$ 可取遍 $\mathfrak{R}_n^*$），故若再写 $\mathscr{T}\boldsymbol{a}=\mathscr{T}\tilde{\boldsymbol{a}}$（$\mathscr{T}$ 取遍 $\mathfrak{R}_n^*$），则 $\tilde{\boldsymbol{a}}\in\mathfrak{R}_n^{**}$，而 $\boldsymbol{a}\to\tilde{\boldsymbol{a}}$ 为 $\mathfrak{R}_n$ 到 $\mathfrak{R}_n^{**}$ 内的映射．又，映射 $\boldsymbol{a}\to\tilde{\boldsymbol{a}}$ 为 $\mathfrak{R}_n$ 到 $\mathfrak{R}_n^{**}$ 上的同构映射（这是因为，$\boldsymbol{a}+\boldsymbol{b}\to\mathscr{T}(\boldsymbol{a}+\boldsymbol{b})=(\boldsymbol{a}+\boldsymbol{b})\mathscr{T}=\boldsymbol{a}\mathscr{T}+\boldsymbol{b}\mathscr{T}=\mathscr{T}\boldsymbol{a}+\mathscr{T}\boldsymbol{b}=\mathscr{T}\tilde{\boldsymbol{a}}+\mathscr{T}\tilde{\boldsymbol{b}}$, $\lambda\boldsymbol{a}\to\mathscr{T}(\lambda\boldsymbol{a})=(\lambda\boldsymbol{a})\mathscr{T}=\lambda(\boldsymbol{a}\mathscr{T})=\lambda(\mathscr{T}\boldsymbol{a})=\lambda\cdot\mathscr{T}\tilde{\boldsymbol{a}}$，并在上面已证明过 $\boldsymbol{a}\to\tilde{\boldsymbol{a}}$ 为 1-1 对应的，于是由于 $\mathfrak{R}_n$ 及 $\mathfrak{R}_n^{**}$ 之维数均等于 n，故 $\mathfrak{R}_n\cong\mathfrak{R}_n^{**}$）．正是由于这样的原因，我们可把 $\mathfrak{R}_n$ 与 $\mathfrak{R}_n^*$ 之共轭空间 $\mathfrak{R}_n^{**}$ 看作是一致的．利用这一点，可证下面的

定理 4 若$(\boldsymbol{e}_1^*, \boldsymbol{e}_2^*, \cdots, \boldsymbol{e}_n^*)$为 $\mathfrak{R}_n^*$ 之一组基底，则必有 $\mathfrak{R}_n$ 之一组基底$(\boldsymbol{e}_1, \boldsymbol{e}_2, \cdots, \boldsymbol{e}_n)$，使得 $\boldsymbol{e}_i\boldsymbol{e}_j^*=\delta_{ij}$（Kronecker 符号）．

证明 我们知道，在 $\mathfrak{R}_n^*$ 之共轭空间 $\mathfrak{R}_n^{**}$ 内能找得一基底$(\boldsymbol{e}_1^{**}, \boldsymbol{e}_2^{**}, \cdots, \boldsymbol{e}_n^{**})$使为$(\boldsymbol{e}_1^*, \boldsymbol{e}_2^*, \cdots, \boldsymbol{e}_n^*)$之补基底，即 $\boldsymbol{e}_i^*\boldsymbol{e}_j^{**}=\delta_{ij}$．这样，再利用上述形成 $\mathfrak{R}_n\cong\mathfrak{R}_n^{**}$ 之关系，有 $\mathfrak{R}_n$ 中 n 个元$(\boldsymbol{e}_1, \boldsymbol{e}_2, \cdots, \boldsymbol{e}_n)$使 $\boldsymbol{e}_j\to\tilde{\boldsymbol{e}}_j=\boldsymbol{e}_j^{**}$，这是说：对任 $\mathscr{T}\in\mathfrak{R}_n^*$，恒有 $\boldsymbol{e}_j\mathscr{T}=\mathscr{T}\boldsymbol{e}_j=\mathscr{T}\tilde{\boldsymbol{e}}_j=\mathscr{T}\boldsymbol{e}_j^{**}$，于是 $\boldsymbol{e}_j\boldsymbol{e}_i^*=\boldsymbol{e}_i^*\boldsymbol{e}_j^{**}=\delta_{ij}$．至于$(\boldsymbol{e}_1, \boldsymbol{e}_2, \cdots, \boldsymbol{e}_n)$为 $\mathfrak{R}_n$ 之基底由 $\mathfrak{R}_n\cong\mathfrak{R}_n^{**}$ 是显然的，或可直接证明于下：设

$$\sum_{j=1}^{n}k_j\boldsymbol{e}_j=\boldsymbol{0},$$

则

$$0=\boldsymbol{0}\boldsymbol{e}_i^*=\sum_{j=1}^{n}k_j(\boldsymbol{e}_j\boldsymbol{e}_i^*)=\sum_{j=1}^{n}k_j\delta_{ij}=k_i,$$

即 $\boldsymbol{e}_1, \cdots, \boldsymbol{e}_n$ 线性无关，故为 $\mathfrak{R}_n$ 之一组基底．

现在再利用共轭空间 $\mathfrak{R}_n^*$ 的概念，谈谈 $\mathfrak{R}_n$ 之某线性变换所对应之矩阵的转置矩阵究竟表示怎样的线性变换．

设 $\mathfrak{R}_n$ 之一组基底为$(\boldsymbol{e}_1, \boldsymbol{e}_2, \cdots, \boldsymbol{e}_n)$，令$(\boldsymbol{e}_1^*, \boldsymbol{e}_2^*, \cdots, \boldsymbol{e}_n^*)$为相对于$(\boldsymbol{e}_1, \boldsymbol{e}_2, \cdots, \boldsymbol{e}_n)$之 $\mathfrak{R}_n^*$ 的补基底，于是 $\boldsymbol{e}_i\boldsymbol{e}_j^*=\delta_{ij}$．令

后，所谓 $\mathfrak{R}_n$ 之某线性变换所对应之矩阵，指的是对基底 $(\boldsymbol{e}_1, \boldsymbol{e}_2, \cdots, \boldsymbol{e}_n)$ 而言，$\mathfrak{R}_n^*$ 之线性变换对应的矩阵则是指对基底 $(\boldsymbol{e}_1^*, \boldsymbol{e}_2^*, \cdots, \boldsymbol{e}_n^*)$ 而言的．令 $\mathscr{L}$ 表示 $\mathfrak{R}_n$ 之一线性变换，其所对应之矩阵为 $\boldsymbol{A}=(a_{ij})$，即 $\boldsymbol{e}_i\mathscr{L}=\sum\limits_{p=1}^{n}a_{ip}\boldsymbol{e}_p$．若 $\mathscr{T}\in\mathfrak{R}_n^*$，即 $\mathscr{T}$ 为 $\mathfrak{R}_n$ 到 $\mathfrak{C}$ 内之线性变换，则易知 $\mathscr{L}\mathscr{T}$ 亦为 $\mathfrak{R}_n$ 到 $\mathfrak{C}$ 内的线性变换，即 $\mathscr{L}\mathscr{T}\in\mathfrak{R}_n^*$．让 $\mathscr{T}$ 取遍 $\mathfrak{R}_n^*$ 时，不难验证映射 $\mathscr{T}\to\mathscr{L}\mathscr{T}$ 是 $\mathfrak{R}_n^*$ 的一个线性变换(这是因为 $\mathscr{L}(\mathscr{T}+\mathscr{S})=\mathscr{L}\mathscr{T}+\mathscr{L}\mathscr{S}$ 及 $\mathscr{L}(\mathscr{T}k)=(\mathscr{L}\mathscr{T})k$)．这个线性变换($\mathfrak{R}_n^*$ 的)是由 $\mathfrak{R}_n$ 之线性变换 $\mathscr{L}$ 所引起的，因而通常用符号 $\mathscr{L}^*$ 表示，叫做 $\mathscr{L}$ 的共轭变换．因而 $\mathscr{T}\mathscr{L}^*=\mathscr{L}\mathscr{T}$，$(\mathscr{T}\in\mathfrak{R}_n^*)$．令 $\mathscr{L}^*$ 对应之矩阵为 $\boldsymbol{B}=(b_{ij})$，即

$$\boldsymbol{e}_i^*\mathscr{L}^*=\sum_{q=1}^{n}b_{iq}\boldsymbol{e}_q^*.$$

于是，因

$$(\boldsymbol{e}_i\mathscr{L})\boldsymbol{e}_j^*=\boldsymbol{e}_i(\mathscr{L}\boldsymbol{e}_j^*)=\boldsymbol{e}_i(\boldsymbol{e}_j^*\mathscr{L}^*),$$

故再以 $\boldsymbol{e}_i\mathscr{L}=\sum\limits_{p}a_{ip}\boldsymbol{e}_p$ 代入左端以及以 $\boldsymbol{e}_j^*\mathscr{L}^*=\sum\limits_{q}b_{jq}\boldsymbol{e}_q^*$ 代入右端，分别可知：

$$(\boldsymbol{e}_i\mathscr{L})\boldsymbol{e}_j^*=(\sum_p a_{ip}\boldsymbol{e}_p)\boldsymbol{e}_j^*=\sum_p a_{ip}(\boldsymbol{e}_p\boldsymbol{e}_j^*)=\sum_p a_{ip}\delta_{pj}=a_{ij},$$

$$\boldsymbol{e}_i(\boldsymbol{e}_j^*\mathscr{L}^*)=\boldsymbol{e}_i(\sum_q b_{jq}\boldsymbol{e}_q^*)=\sum_q b_{jq}(\boldsymbol{e}_i\boldsymbol{e}_q^*)=\sum_q b_{jq}\delta_{iq}=b_{ji};$$

于是有 $b_{ji}=a_{ij}$，即 $\boldsymbol{B}=\boldsymbol{A}'$．这就是说：若 $\mathfrak{R}_n$ 之线性变换 $\mathscr{L}$ 关于基底 $(\boldsymbol{e}_1, \boldsymbol{e}_2, \cdots, \boldsymbol{e}_n)$ 之矩阵为 $\boldsymbol{A}$，那末由 $\mathscr{L}$ 所引起的共轭空间 $\mathfrak{R}_n^*$ 中的线性变换 $\mathscr{L}^*$(即 $\mathscr{L}$ 之共轭变换)关于补基底 $(\boldsymbol{e}_1^*, \boldsymbol{e}_2^*, \cdots, \boldsymbol{e}_n^*)$ 之矩阵就等于 $\boldsymbol{A}$ 之转置矩阵．

又，若 β 为 $\mathfrak{R}_n$ 之线性变换，则因对每 $\mathscr{T}\in\mathfrak{R}_n^*$ 恒有

$$\begin{aligned}\mathscr{T}(\alpha+\beta)^*&=(\alpha+\beta)\mathscr{T}=\alpha\mathscr{T}+\beta\mathscr{T}\\&=\mathscr{T}\alpha^*+\mathscr{T}\beta^*=\mathscr{T}(\alpha^*+\beta^*),\end{aligned}$$

以及

$$\mathscr{T}(\alpha\beta)^{*}=(\alpha\beta)\mathscr{T}=\alpha(\beta\mathscr{T})=\alpha(\mathscr{T}\beta^{*})$$
$$=(\mathscr{T}\beta^{*})\alpha^{*}=\mathscr{T}(\beta^{*}\alpha^{*}),$$

故有

$$(\alpha+\beta)^{*}=\alpha^{*}+\beta^{*}\quad 及\quad (\alpha\beta)^{*}=\beta^{*}\alpha^{*},$$

因而矩阵之和的转置矩阵等于各个矩阵之转置的和，而矩阵之积的转置就等于各个转置之逆序的积．这也就是说从线性变换的角度再度证明了定理3.

练习十八

1．证明：矩阵之转置符号“′”伴随符号“*”，以及逆的符号“－1”，三者的关系有：

(i) $(\boldsymbol{A}^{-1})'=(\boldsymbol{A}')^{-1}$，(ii) $(\boldsymbol{A}^{-1})^{*}=(\boldsymbol{A}^{*})^{-1}$，(iii) $(\boldsymbol{A}^{*})'=(\boldsymbol{A}')^{*}$.

即对矩阵施行其中任两个符号时，其结果与两个符号的先后顺序无关．

2．设 $\boldsymbol{A}=(a_{ij})$ 为 $n\times(n+1)$-矩阵，$\boldsymbol{E}$ 为 n 级单位矩阵．证明：适合 $\boldsymbol{AX}=\boldsymbol{E}$ 的 $(n+1)\times n$-矩阵 $\boldsymbol{X}$ 存在的充要条件是 $\boldsymbol{A}$ 的秩 $r_{\boldsymbol{A}}=n$.

3．设 $\boldsymbol{A}=(a_{ij})$ 与 $\boldsymbol{B}=(b_{ij})$ 是两个 n 级矩阵，并令

$$\boldsymbol{A}^{(k)}=\begin{pmatrix}a_{11} & \cdots & a_{1n} & b_{1k}\\ \vdots & & \vdots & \vdots\\ a_{n1} & \cdots & a_{nn} & b_{nk}\end{pmatrix},\quad (k=1,\ 2,\ \cdots,\ n).$$

证明有 n 级矩阵 $\boldsymbol{X}$ 能使 $\boldsymbol{AX}=\boldsymbol{B}$ 成立的充要条件是 $\boldsymbol{A}$，$\boldsymbol{A}^{(1)}$，$\boldsymbol{A}^{(2)}$，…，$\boldsymbol{A}^{(n)}$ 都有相等的秩．

4．设 $\boldsymbol{D}=\begin{pmatrix}\lambda_1\boldsymbol{E}_1 & & & \\ & \lambda_2\boldsymbol{E}_2 & & \\ & & \ddots & \\ & & & \lambda_k\boldsymbol{E}_k\end{pmatrix}$，$\boldsymbol{E}_i$ 为 n_i 级单位矩阵．试证：与 $\boldsymbol{D}$ 相乘为可交换的矩阵，必为且仅为

$$\boldsymbol{X}=\begin{pmatrix}\boldsymbol{C}_1 & & & \\ & \boldsymbol{C}_2 & & \\ & & \ddots & \\ & & & \boldsymbol{C}_k\end{pmatrix}$$

形，式中 $\boldsymbol{C}_i$ 为 n_i 级矩阵．但 $\lambda_i\neq\lambda_j$（$i\neq j$ 时）．

（注意，此题为练习十六第1题的推广，这个问题很重要，以后要用．）

5. 设 $\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C}, \boldsymbol{D}$ 都是 n 级矩阵. 若 $\boldsymbol{AC}=\boldsymbol{CA}$, 试证

$$\begin{vmatrix} \boldsymbol{A} & \boldsymbol{B} \\ \boldsymbol{C} & \boldsymbol{D} \end{vmatrix} = |\boldsymbol{AD}-\boldsymbol{CB}|.$$

第 8 节 矩阵乘积的秩

我们知道,矩阵的秩不仅是一个代数的概念,它有很重要的几何意义,即矩阵所对应的线性变换把全空间 $\mathfrak{R}_n$ 映射成的子空间 $\mathfrak{R}$ 的维数就等于这矩阵的秩. 故研究矩阵的秩是一个很有意义的问题.

我们知道, 两个 n 级矩阵之积为满秩的充要条件是各个矩阵都是满秩的. 但它们的秩之间究竟有些怎样的关系呢? 一般有下面的

定理 1 两个矩阵之积的秩, 不能大于这两个矩阵中的任何一个的秩(这里指的一般都是长方矩阵).

证明 设 $\boldsymbol{A}=(a_{ij})$ 及 $\boldsymbol{B}=(b_{ij})$ 分别为 $m\times k$-及 $k\times n$-矩阵, 于是 $\boldsymbol{P}=\boldsymbol{AB}=(p_{ij})$ 是 $m\times n$-矩阵. 由矩阵乘法的规则, $\boldsymbol{P}$ 之每个行向量(n 元向量)是 $\boldsymbol{B}$ 之行向量(n 元的)的线性组合, 因而也必是 $\boldsymbol{B}$ 之行向量里线性无关最大数目的线性组合, 故据第二章第 1 节的定理 2 可知它们的秩满足 $r_P\leqslant r_B$. 同样, 由于 $\boldsymbol{P}$ 之每个列向量(m 元的)是 $\boldsymbol{A}$ 之列向量(m 元的)的线性组合, 又能推知 $r_P\leqslant r_A$. 证完.

当然, 还可以用线性变换的理论来证明定理 1, 证明如下:

设 $\mathscr{T}$ 是从 $\mathfrak{R}_m$ 到 $\mathfrak{R}'_n$ 的线性变换, $\boldsymbol{A}$ 为 $\mathscr{T}$ 之矩阵(关于基底 $\boldsymbol{\alpha}_1, \cdots, \boldsymbol{\alpha}_m$ 与 $\boldsymbol{\alpha}'_1, \cdots, \boldsymbol{\alpha}'_n$), $\mathscr{S}$ 是从 $\mathfrak{R}'_n$ 到 $\mathfrak{R}''_k$ 的线性变换, $\boldsymbol{B}$ 为 $\mathscr{S}$ 之矩阵(关于基底 $\boldsymbol{\alpha}'_1, \cdots, \boldsymbol{\alpha}'_n$ 与 $\boldsymbol{\alpha}''_1, \cdots, \boldsymbol{\alpha}''_k$). 于

是 r_A 等于由 $\boldsymbol{\alpha}_i\mathscr{T}(i=1, \cdots, m)$生成的 $\mathfrak{R}'_n$ 之子空间 $\mathfrak{R}'_n$ 的维数，而 r_B 等于由 $\boldsymbol{\alpha}'_i\mathscr{S}(i=1, \cdots, n)$生成的 $\mathfrak{R}''_k$ 之子空间 $\mathfrak{R}''_k$ 的维数. 故 $\mathfrak{R}_m\mathscr{T}=\mathfrak{R}'_n\sqsubseteq\mathfrak{R}'_n$，$\mathfrak{R}'_n\mathscr{S}=\mathfrak{R}''_k\sqsubseteq\mathfrak{R}''_k$，于是有

$$\mathfrak{R}_m(\mathscr{T}\mathscr{S})=(\mathfrak{R}_m\mathscr{T})\mathscr{S}=\mathfrak{R}'_n\mathscr{S}\sqsubseteq\mathfrak{R}'_n\mathscr{S}=\mathfrak{R}''_k,$$

即 $\mathfrak{R}_m(\mathscr{T}\mathscr{S})$ 之维数 $\leqslant\mathfrak{R}''_k$ 之维数；但 $\mathfrak{R}_m(\mathscr{T}\mathscr{S})$ 之维数等于 $\boldsymbol{P}=\boldsymbol{AB}$ 之秩 r_{AB}，而 $\mathfrak{R}''_k$ 之维数等于 $\boldsymbol{B}$ 之秩 r_B，因而有 $r_{AB}\leqslant r_B$，定理 1 得证.

如 $\boldsymbol{A}$、$\boldsymbol{B}$ 中有一为方阵，例如 $\boldsymbol{A}$ 是方阵，则当 $\boldsymbol{A}$ 为满秩时，由 $\boldsymbol{P}=\boldsymbol{AB}$ 可得 $\boldsymbol{B}=\boldsymbol{A}^{-1}\boldsymbol{P}$，故据定理 1，有 $r_B\leqslant r_P$. 但由等式 $\boldsymbol{P}=\boldsymbol{AB}$ 再一次运用定理 1，可得 $r_P\leqslant r_B$. 故有 $r_P=r_B$. 同理，若 $\boldsymbol{P}=\boldsymbol{AB}$ 中 $\boldsymbol{B}$ 为满秩时，又必有 $r_P=r_A$. 于是证得了：

推论 一个矩阵，不论在它的左边或右边用任何满秩方阵去乘，不改变它的秩.

这推论还可直接证明于下：例如，当 $\boldsymbol{A}$ 为满秩时，则 $\boldsymbol{A}$ 为有限多个初等矩阵之积，因之 $\boldsymbol{AB}$ 可看作是由 $\boldsymbol{B}$ 之行施行若干回初等变换而成，即 $\boldsymbol{AB}\simeq\boldsymbol{B}$，故 $r_{AB}=r_B$.

若想看出积 $\boldsymbol{P}=\boldsymbol{AB}$ 之 r 级子式与 $\boldsymbol{A}$ 或 $\boldsymbol{B}$ 之 r 级子式的关系，还有较深刻的结果. 先引进必要的记号：即从矩阵 $\boldsymbol{A}$ 中取其第 $i_1, i_2, \cdots, i_r$ 行及第 $j_1, j_2, \cdots, j_r$ 列 $(i_1<i_2<\cdots<i_r, j_1<j_2<\cdots<j_r)$ 组成的 r 级子式用符号 $d(\boldsymbol{A}^{i_1i_2\cdots i_r}_{j_1j_2\cdots j_r})$ 记之，即

$$d(\boldsymbol{A}^{i_1i_2\cdots i_r}_{j_1j_2\cdots j_r})=\begin{vmatrix} a_{i_1j_1} & a_{i_1j_2} & \cdots & a_{i_1j_r}\\ a_{i_2j_1} & a_{i_2j_2} & \cdots & a_{i_2j_r}\\ \vdots & \vdots & & \vdots\\ a_{i_rj_1} & a_{i_rj_2} & \cdots & a_{i_rj_r}\end{vmatrix}.$$

定理 2 设 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 分别为 $m\times k$–及 $k\times n$–矩阵，$\boldsymbol{P}=$

$\boldsymbol{AB}$，则

$$d(\boldsymbol{P}^{i_1i_2\cdots i_r}_{j_1j_2\cdots j_r})=\sum_{(h_1,\cdots,h_r)}d(\boldsymbol{A}^{i_1i_2\cdots i_r}_{h_1h_2\cdots h_r})\cdot d(\boldsymbol{B}^{h_1h_2\cdots h_r}_{j_1j_2\cdots j_r}),$$

其中 $\sum\limits_{(h_1,\cdots,h_r)}$ 表示从 k 个自然数 1，2，…，k 中取 r 个的所有组合形成的和。

证明

$$d(\boldsymbol{P}^{i_1i_2\cdots i_r}_{j_1j_2\cdots j_r})=\begin{vmatrix}p_{i_1j_1} & p_{i_1j_2} & \cdots & p_{i_1j_r}\\ p_{i_2j_1} & p_{i_2j_2} & \cdots & p_{i_2j_r}\\ \vdots & \vdots & & \vdots\\ p_{i_rj_1} & p_{i_rj_2} & \cdots & p_{i_rj_r}\end{vmatrix}$$

$$=\begin{vmatrix}\sum\limits_{t_1=1}^{k}a_{i_1t_1}b_{t_1j_1} & \sum\limits_{t_2=1}^{k}a_{i_1t_2}b_{t_2j_2} & \cdots & \sum\limits_{t_r=1}^{k}a_{i_1t_r}b_{t_rj_r}\\ \vdots & \vdots & & \vdots\\ \sum\limits_{t_1=1}^{k}a_{i_rt_1}b_{t_1j_1} & \sum\limits_{t_2=1}^{k}a_{i_rt_2}b_{t_2j_2} & \cdots & \sum\limits_{t_r=1}^{k}a_{i_rt_r}b_{t_rj_r}\end{vmatrix}$$

$$=\sum_{t_1,\cdots,t_r=1}^{k}\begin{vmatrix}a_{i_1t_1} & a_{i_1t_2} & \cdots & a_{i_1t_r}\\ a_{i_2t_1} & a_{i_2t_2} & \cdots & a_{i_2t_r}\\ \vdots & \vdots & & \vdots\\ a_{i_rt_1} & a_{i_rt_2} & \cdots & a_{i_rt_r}\end{vmatrix}b_{t_1j_1}b_{t_2j_2}\cdots b_{t_rj_r},$$

而在和 $\sum\limits_{t_1,\cdots,t_r=1}^{k}$ 的 k^r 个项中凡有两个 $t_i(i=1,\cdots,r)$ 相等的项显然等于零(因这时行列式有两列相同)，因此只考虑其中所有的 t 为互异的情况．于是，

$$d(\boldsymbol{P}^{i_1i_2\cdots i_r}_{j_1j_2\cdots j_r})=\sum_{\substack{1\leq t_1,t_2,\cdots,t_r\leq k\\ t_i\neq t_j(i\neq j\text{时})}}\begin{vmatrix}a_{i_1t_1} & a_{i_1t_2} & \cdots & a_{i_1t_r}\\ a_{i_2t_1} & a_{i_2t_2} & \cdots & a_{i_2t_r}\\ \vdots & \vdots & & \vdots\\ a_{i_rt_1} & a_{i_rt_2} & \cdots & a_{i_rt_r}\end{vmatrix}b_{t_1j_1}b_{t_2j_2}\cdots b_{t_rj_r}. \tag{23}$$

因为 $t_1, t_2, \cdots, t_r$ 中无两个相等，故 $t_1, t_2, \cdots, t_r$ 是从

1, 2, …, k 中取 r 个的一个组合；今令 $t_1, t_2, \cdots, t_r$ 是 $h_1, h_2, \cdots, h_r (h_1<h_2<\cdots<h_r)$ 的一个排列，则(23)中之 r 级行列式就等于

$$(-1)^{[t_1, t_2, \cdots, t_r]} d(A_{h_1 h_2 \cdots h_r}^{i_1 i_2 \cdots i_r}) \tag{24}$$

(参看第一章第 2 节的例 1)．故若将(23)式右边和 $\sum$ 内凡由 $h_1, h_2, \cdots, h_r$ 之排列 $t_1, t_2, \cdots, t_r$ (共有 $r!$ 个这样的排列)形成的 r 级行列式都用(24)去替换，就得知：

$$d(P_{j_1 j_2 \cdots j_r}^{i_1 i_2 \cdots i_r}) = \sum_{1 \le h_1 < \cdots < h_r \le k} [d(A_{h_1 h_2 \cdots h_r}^{i_1 i_2 \cdots i_r}) \cdot \sum_{(t_1 \cdots t_r)}' (-1)^{[t_1, t_2, \cdots, t_r]} b_{t_1 j_1} b_{t_2 j_2} \cdots b_{t_r j_r}], \tag{25}$$

但 $\sum\limits_{1 \le h_1 < \cdots < h_r \le k}$ 表示从 1, 2, …, k 这 k 个数内任取 r 个数 $h_1, h_2, \cdots, h_r$ 使 $h_1<h_2<\cdots<h_r$ 的一切组合，而 $\sum\limits_{(t_1 \cdots t_r)}'$ 是表示就某一组固定的 $h_1, h_2, \cdots, h_r (h_1<h_2<\cdots<h_r)$ 上的一切排列所成之诸项的和．于是，据行列式之定义，知

$$\sum_{(t_1 \cdots t_r)}' (-1)^{[t_1, \cdots, t_r]} b_{t_1 j_1} b_{t_2 j_2} \cdots b_{t_r j_r} = d(B_{j_1 j_2 \cdots j_r}^{h_1 h_2 \cdots h_r}),$$

将上式代入(25)式内，就得到

$$d(P_{j_1 j_2 \cdots j_r}^{i_1 i_2 \cdots i_r}) = \sum_{1 \le h_1 < \cdots < h_r \le k} d(A_{h_1 h_2 \cdots h_r}^{i_1 i_2 \cdots i_r}) \cdot d(B_{j_1 j_2 \cdots j_r}^{h_1 h_2 \cdots h_r}),$$

定理 2 证完．

特别地，当 $m=k=n$ 时，A、B、P 都是 n 级矩阵．如又 $r=n$，则 $\sum\limits_{1 \le h_1 < \cdots < h_n \le n}$ 内只有一项，因而 $h_1=1, \cdots, h_n=n$，并且 $i_1=j_1=1, \cdots, i_n=j_n=n$，这时定理 2 就成了 $|P|=|A|\cdot|B|$，这就是说，第 131 页的定理是这定理 2 的一个特例．

由定理 2，即可得到下面的

推论 设 A、B 分别为 $m\times k$-及 $k\times m$-矩阵，$P=AB$ 当然是 m 级矩阵．若 $m>k$，则 P 之行列式为 0，即 $|P|=0$．

证明 令定理 2 中的 $r=m$，则从 1, 2, …, k 中取 m 个的任意组合里都至少有两个数相等，因而定理 2 中右边和 $\sum$ 内每一项都等于零，这就证得了 $|\boldsymbol{P}|=0$.

再考虑定理 2 中矩阵 $\boldsymbol{A}$ 或 $\boldsymbol{B}$ 的秩小于 r 的情形. 这时，据定理 2 可知 $\boldsymbol{P}=\boldsymbol{AB}$ 中一切 r 级子式必为零，故若 $\boldsymbol{A}$ (或 $\boldsymbol{B}$) 的秩 $=r_1$，则取 $r=r_1+1$，即知 $r_P\leqslant r_1$，这又证明了定理 1，即可把定理 1 看做是定理 2 的一个推论.

定理 2 的用处很广，可推得一个重要的不等式，即柯西 (Cauchy) 不等式：假定

$$\boldsymbol{A}=\begin{pmatrix} x_1 & x_2 & \cdots & x_n \\ y_1 & y_2 & \cdots & y_n \end{pmatrix},$$

则

$$\boldsymbol{P}=\boldsymbol{AA}'=\begin{pmatrix} \sum\limits_{i=1}^{n} x_i^2 & \sum\limits_{i=1}^{n} x_i y_i \\ \sum\limits_{i=1}^{n} x_i y_i & \sum\limits_{i=1}^{n} y_i^2 \end{pmatrix},$$

因此，$\boldsymbol{P}$ 的行列式

$$|\boldsymbol{P}|=\left(\sum_{i=1}^{n} x_i^2\right)\left(\sum_{i=1}^{n} y_i^2\right)-\left(\sum_{i=1}^{n} x_i y_i\right)^2.$$

但由定理 2，又有

$$|\boldsymbol{P}|=\sum_{1\leqslant i<j\leqslant n}\begin{vmatrix} x_i & x_j \\ y_i & y_j \end{vmatrix}\cdot\begin{vmatrix} x_i & y_i \\ x_j & y_j \end{vmatrix}=\sum_{1\leqslant i<j\leqslant n}(x_i y_j-x_j y_i)^2,$$

故当 x_i 与 y_j 都为实数时，$|\boldsymbol{P}|\geqslant 0$ (等于 0 的情形是在且仅在 $x_1, x_2, \cdots, x_n$ 与 $y_1, y_2, \cdots, y_n$ 成比例时)，因而有

$$\left(\sum_{i=1}^{n} x_i^2\right)\left(\sum_{i=1}^{n} y_i^2\right)\geqslant\left(\sum_{i=1}^{n} x_i y_i\right)^2.$$

这就是柯西不等式.

[例 1] 在秩为 r 的 $m\times n$-矩阵 $\boldsymbol{A}$ 中，任取 s 个行，得到 $s\times n$-矩阵 $\boldsymbol{B}$. 证明：$\boldsymbol{B}$ 之秩不小于 $r+s-m$. 同理，取 $\boldsymbol{A}$

之任 s 个列，得到 $m\times s$-矩阵 $\boldsymbol{C}$，$\boldsymbol{C}$ 之秩 $\geqslant r+s-n$.

证明 因 $\boldsymbol{A}$ 中必有 r 个行是线性无关的，而剩下 $m-r$ 个行在 $s>m-r$ 时即令全部都出现在 $\boldsymbol{B}$ 内，这时 $\boldsymbol{B}$ 也要包含上述 r 个行中的 $s-(m-r)=r+s-m$ 个行；若所剩下 $m-r$ 个行不全部出现在 $\boldsymbol{B}$ 内，那末 $\boldsymbol{B}$ 至少要包含上述 r 个行的 $r+s-m$ 个. 故 $\boldsymbol{B}$ 中至少有 $r+s-m$ 个行是线性无关的，即 $r_B\geqslant r+s-m$. 这说明了：当 $s>m-r$ 时论断是成立的. 至于在 $s\leqslant m-r$ 时，$r+s-m\leqslant 0$，论断当然也成立.

［例 2］ 设 $\boldsymbol{A}$，$\boldsymbol{B}$ 是两个 n 级矩阵，证明：

$$r_A+r_B-n\leqslant r_{AB}\leqslant \min(r_A,\ r_B).$$

证明 由第二章第 4 节可知 $\boldsymbol{A}\simeq\boldsymbol{E}_0$，但 $\boldsymbol{E}_0$ 为从 n 级单位矩阵 $\boldsymbol{E}$ 中删去主对角线上后面的一些 1 而只保留前 r_A 个 1 所成的 n 级矩阵. 再由第 5 节定理 6 的推论，知有两个 n 级满秩矩阵 $\boldsymbol{P}$ 与 $\boldsymbol{Q}$，使 $\boldsymbol{A}=\boldsymbol{P}\boldsymbol{E}_0\boldsymbol{Q}$；于是

$$\boldsymbol{AB}=\boldsymbol{PE}_0\boldsymbol{QB}=\boldsymbol{PE}_0\boldsymbol{C},\ (\boldsymbol{C}=\boldsymbol{QB})$$

故 $r_C=r_B$. 若令 $\boldsymbol{C}=(c_{ij})$，则易验证

$$\boldsymbol{E}_0\boldsymbol{C}=\begin{pmatrix} c_{11} & c_{12} & \cdots & c_{1n}\\ \vdots & \vdots & & \vdots\\ c_{r_A1} & c_{r_A2} & \cdots & c_{r_An}\\ 0 & 0 & \cdots & 0\\ \vdots & \vdots & & \vdots\\ 0 & 0 & \cdots & 0\end{pmatrix},$$

因而据例 1 知

$$r_{E_0C}\geqslant r_C+r_A-n=r_B+r_A-n.$$

再由 $\boldsymbol{P}$ 之满秩性，可知

$$r_{AB}=r_{PE_0C}=r_{E_0C}\geqslant r_A+r_B-n.$$

至于 $r_{AB}\leqslant\min(r_A,\ r_B)$，根据定理 1 是显然的. 证完.

练 习 十 九

1．证明：凡秩为 r 的矩阵，总可以写为 r 个秩为 1 的矩阵之和。

2．设 $\boldsymbol{A}_1, \boldsymbol{A}_2, \cdots, \boldsymbol{A}_k$ 都是 n 级矩阵(降、满秩均可)，试证：

$$(\boldsymbol{A}_1\boldsymbol{A}_2\cdots\boldsymbol{A}_k)^*=\boldsymbol{A}_k^*\cdots\boldsymbol{A}_2^*\boldsymbol{A}_1^*,$$

其中 $\boldsymbol{A}^*$ 表示 $\boldsymbol{A}$ 的伴随矩阵。

第五章 λ-矩 阵

前几章所讨论的矩阵，矩阵的元素都是通常的数，但在把许多实际问题化为数学问题来处理时，往往会遇到以某个变量的函数（连续的或有时不连续的）做元素的矩阵．本章研究以一个变量的多项式做元素的矩阵，习惯上叫这样的矩阵为λ-矩阵，即以变量λ的多项式做元素的矩阵．

在第二章，我们已经知道，矩阵的秩在研究矩阵时占重要地位，影响一个矩阵的秩或者说与秩紧密地联系着的，是初等变换、等价以及标准形等问题．本章就λ-矩阵来研究这些问题，具体说，讨论下面几个问题：

1．什么是λ-矩阵的秩？

2．什么是λ-矩阵的初等变换？

3．怎样的两个λ-矩阵是等价的？怎样判定？

4．什么是λ-矩阵的标准形？

上面这些问题都集中在第1～2节内解决．在第3节里讨论一般多项式的哪些性质可推广到λ-矩阵．从第4节以后就讨论怎样利用λ-矩阵来研究以数为元素的矩阵．

第1节 λ-矩阵的基本概念

这里，首先研究λ-矩阵的秩．我们知道，λ-矩阵中任意级的子行列式一般是λ的多项式，当变量λ取适当的值时，总可以使它等于零，因此，λ-矩阵的秩就不能象前面元素是数时

那样从它的子行列式可否取零值(当 λ 为适当的数值时)来确定. 然而, λ 的多项式有为恒等于零的(即当 λ 取任何数值时这多项式的值总是等于零), 也有不恒等于零的. 于是, 可以根据是否恒等于零来研究 λ-矩阵的秩.

定义 1 若 λ-矩阵中不恒等于零之子行列式的最大级数是 r, 就叫这 λ-矩阵的秩为 r. 当 λ-矩阵是方阵时,根据它的行列式恒等于零或不恒等于零,分别叫它为**降秩的**或**满秩的**.

为要研究如何确定一个 λ-矩阵的秩, 就需要知道在什么时候 λ-矩阵的秩不改变, 这就自然而然地联系到类似于第二、四章里所谈的初等变换了.

同过去讲的一样,容易知道: 互换 λ-矩阵的任两行(列), 或用不等于零的数 a 乘任一行(列), 或用一个 λ 的多项式 $f(\lambda)$ 乘某行(列)后再加到另一行(列)上去, 其结果都不改变 λ-矩阵的秩, 因此, 这样一些变换在研究 λ-矩阵时也同样重要,故特给出下面的

定义 2 下列各种类型的变换,叫做 λ-矩阵的**初等变换**:

1° 任二行(列)的互换,

2° 用数 $a(\neq 0)$乘任一行(列),

3° 用 λ 的一多项式乘某行(列)后再加到另一行(列)上去.

又,同第二章一样, 也可给出两个 λ-矩阵等价的概念,即

定义 3 一个 λ-矩阵 $\boldsymbol{A}(\lambda)$ 如经过有限次初等变换后变到另一个 λ-矩阵 $\boldsymbol{B}(\lambda)$,就叫 $\boldsymbol{A}(\lambda)$ 与 $\boldsymbol{B}(\lambda)$ **等价**,用符号 $\boldsymbol{A}(\lambda) \simeq \boldsymbol{B}(\lambda)$ 表示.

关于定义 2, 作如下一点补充说明: 对比第二章第 4 节中的定义, 可发现 1°、3° 两类型是完全相仿的, 即可平行地叙述,而 2° 型却不相仿. 试问: 定义 2 中,为什么不用一个不恒

等于零的 λ 之多项式 $f(\lambda)$ 去乘，而是要用一个数 $\alpha(\neq 0)$ 去乘呢？

固然，象这样改了之后，λ-矩阵的秩确是不变，可是，根据定义 2，初等变换都有逆变换（譬如，用交换两行可以从 $\boldsymbol{A}(\lambda)$ 得到 $\boldsymbol{B}(\lambda)$ 时，则用它相反的交换两行就可以从 $\boldsymbol{B}(\lambda)$ 得到 $\boldsymbol{A}(\lambda)$；又如，用 $f(\lambda)$ 乘第 i 行后再加到第 j 行上去可使 $\boldsymbol{A}(\lambda)$ 变为 $\boldsymbol{B}(\lambda)$，则反过来，用 $-f(\lambda)$ 乘第 i 行后再加到第 j 行上去又可使 $\boldsymbol{B}(\lambda)$ 变为 $\boldsymbol{A}(\lambda)$；再，若用数 $\alpha\neq 0$ 乘第 i 行可自 $\boldsymbol{A}(\lambda)$ 变到 $\boldsymbol{B}(\lambda)$，则用数 $\alpha^{-1}\neq 0$ 乘第 i 行就可以自 $\boldsymbol{B}(\lambda)$ 变到 $\boldsymbol{A}(\lambda)$）．并且逆变换也是初等变换，所以"等价"适合

1．自反律：$\boldsymbol{A}(\lambda)\simeq\boldsymbol{A}(\lambda)$，

2．对称律：由 $\boldsymbol{A}(\lambda)\simeq\boldsymbol{B}(\lambda)$ 可得 $\boldsymbol{B}(\lambda)\simeq\boldsymbol{A}(\lambda)$，

3．传递律：由 $\boldsymbol{A}(\lambda)\simeq\boldsymbol{B}(\lambda)$ 及 $\boldsymbol{B}(\lambda)\simeq\boldsymbol{C}(\lambda)$ 得
$$\boldsymbol{A}(\lambda)\simeq\boldsymbol{C}(\lambda).$$

因此，可把所有 $m\times n$-λ 矩阵分成许多不同的类，叫做等价类，使在同一类中的 λ-矩阵都是相互等价的，在不同类中的 λ-矩阵都互不等价．因之，不必对一个一个 λ-矩阵单独地研究，只要把等价类中某一个具有代表性的 λ-矩阵研究清楚了，同一类中所有 λ-矩阵的某些性质（它们只依赖于初等变换）也就清楚了，例如秩相等是其中一个性质．

在定义 2 里，如果第 2° 类型的初等变换中用"一个不恒等于零的 λ 之多项式 $\varphi(\lambda)$"来代替"不为零的数 α"，那末这变换就没有逆变换（譬如，把 $\boldsymbol{A}(\lambda)$ 之第 i 行用 $\varphi(\lambda)(\neq 0)$ 去乘而得到 $\boldsymbol{B}(\lambda)$，则自 $\boldsymbol{B}(\lambda)$ 要得到 $\boldsymbol{A}(\lambda)$ 时就需要用 $\dfrac{1}{\varphi(\lambda)}$ 去乘第 i 行，但 $\dfrac{1}{\varphi(\lambda)}$ 已不再是 λ 的多项式了）．因为这样改了之后的变换没有逆变换，所以对称律不成立，于是如上的分类

也就不可能，也就不能在所谓等价类中讨论问题，这当然不是我们所期望的.

我们已经知道，两个 λ-矩阵若等价，就必有等秩. 可是，它的逆不成立，此与元素是数的矩阵不同(第二章第 4 节的定理 3). 也就是说，秩相等的两个 λ-矩阵不一定就等价，这只要看下面的例子就行了. 例如

$$\boldsymbol{A}(\lambda)=\begin{pmatrix}\lambda & 1\\ 0 & \lambda\end{pmatrix},\quad \boldsymbol{B}(\lambda)=\begin{pmatrix}1 & -\lambda\\ 1 & \lambda\end{pmatrix},$$

它们的行列式分别是 $|\boldsymbol{A}(\lambda)|=\lambda^2$, $|\boldsymbol{B}(\lambda)|=2\lambda$, 显然它们的秩都是 2. 然由初等变换的定义，容易知道两个等价的 λ-矩阵之行列式相差只不过是一个不为零的常数因子. 因此，$\boldsymbol{A}(\lambda)$ 与 $\boldsymbol{B}(\lambda)$ 不等价. 这就是说：秩相等是 λ-矩阵等价的必要条件，但非充分条件. 两个 λ-矩阵等价的必要而又充分的条件是什么呢？下一节就来解答这个问题.

第 2 节　不变因子、初等因子

前节已指出，秩相等是两个 λ-矩阵等价的必要条件，那么，除了秩相等之外，还应该具有什么性质，才能使两个 λ-矩阵等价呢？

当 λ-矩阵 $\boldsymbol{A}(\lambda)$ 的秩 $=r$ 时，它里面凡高于 r 级的子行列式都恒等于零，但却至少有一个 r 级子行列式不恒等于零；而且，由拉普拉斯展开式又知，在 $\boldsymbol{A}(\lambda)$ 中的 $k(\leq r)$ 级子行列式里也至少有一个不恒等于零. 研究 $\boldsymbol{A}(\lambda)$ 中不恒等于零的 i 级 $(i\leqslant r)$ 子行列式(每个不恒等于零的 i 级子行列式是 λ 的多项式)，这样的 λ 的多项式之最高公因式用符号 $D_i^A(\lambda)$ 表示，也可简单地用 $D_i(\lambda)$ 来表示(如果在研究过程中没有突出

λ-矩阵 $A(\lambda)$ 的必要时). 我们知道: 多项式的最高公因式不是唯一的, 彼此相差一个常数因子, 为确定起见, $D_i(\lambda)$ 的最高次项的系数总是规定为1.

今假定 $\boldsymbol{A}(\lambda)\simeq\boldsymbol{B}(\lambda)$. 如果只是利用了前节定义 2 的第 1°、2° 两种类型的初等变换自 $\boldsymbol{A}(\lambda)$ 得到 $\boldsymbol{B}(\lambda)$, 显然就有 $D_i^A(\lambda)\equiv D_i^B(\lambda)$, $(i=1, 2, \cdots)$. 如果自 $\boldsymbol{A}(\lambda)$ 得到 $\boldsymbol{B}(\lambda)$ 的过程中利用过了第 3° 类型的初等变换, 譬如, $\boldsymbol{B}(\lambda)$ 是用多项式 $f(\lambda)$ 乘 $\boldsymbol{A}(\lambda)$ 的第 p 行后再加到第 q 行上去所得到的 λ-矩阵, 则易看出 $\boldsymbol{B}(\lambda)$ 的任一个 i 级子行列式可以写成 $\varphi_i(\lambda)\pm f(\lambda)\cdot\psi_i(\lambda)$ 的形状, 式中 $\varphi_i(\lambda)$ 及 $\psi_i(\lambda)$ 是 $\boldsymbol{A}(\lambda)$ 的两个 i 级子行列式(可能 $\psi_i(\lambda)$ 不出现), 所以 $D_i^A(\lambda)$ 是 $\boldsymbol{B}(\lambda)$ 的任一 i 级子行列式的因式, 故 $D_i^A(\lambda)$ 整除 $D_i^B(\lambda)$, 记成 $D_i^A(\lambda)\mid D_i^B(\lambda)$. 再利用上述变换的逆变换, 可以从 $\boldsymbol{B}(\lambda)$ 得到 $\boldsymbol{A}(\lambda)$, 因此同样又可以得到 $D_i^B(\lambda)\mid D_i^A(\lambda)$. 由于 $D_i^A(\lambda)$ 和 $D_i^B(\lambda)$ 的最高次项的系数都等于 1, 所以 $D_i^A(\lambda)\equiv D_i^B(\lambda)$. 总之, 当 $\boldsymbol{A}(\lambda)\simeq\boldsymbol{B}(\lambda)$ 时, 就有 $D_i^A(\lambda)\equiv D_i^B(\lambda)$ (其中 $i=1, 2, \cdots, r$), 这里 r 是它们的共同秩.

我们现在来证明条件 $D_i^A(\lambda)\equiv D_i^B(\lambda)$ (其中 $i=1, 2, \cdots$) 又是 $\boldsymbol{A}(\lambda)$ 与 $\boldsymbol{B}(\lambda)$ 等价的充分条件. 为了便于说明问题, 先将秩为 r 的 λ-矩阵 $\boldsymbol{A}(\lambda)$ 化成标准形.

若 $r=0$, 则 $\boldsymbol{A}(\lambda)$ 之每元都是 0, 故已不是 λ-矩阵的问题了. 于是可假定 $r\neq0$, 因之 $\boldsymbol{A}(\lambda)$ 里面至少有一元(λ 之多项式)不恒为零. 经过行或列的交换, 可得到一个与 $\boldsymbol{A}(\lambda)$ 等价的 λ-矩阵, 使它左上角的元素不恒等于零. 为简单计, 就令 $\boldsymbol{A}(\lambda)$ 之左上角的元 $a_{11}(\lambda)\not\equiv0$.

若 $\boldsymbol{A}(\lambda)$ 中有元素 $a_{ik}(\lambda)$ 不能被 $a_{11}(\lambda)$ 整除, 假如这样的 $a_{ik}(\lambda)$ 出现在 $\boldsymbol{A}(\lambda)$ 的第一行里面, 例如是 $a_{1k}(\lambda)$, 则 $a_{1k}(\lambda)$ 可

以写成

$$a_{1k}(\lambda)=a_{11}(\lambda)\cdot\varphi(\lambda)+b(\lambda),$$

这里 $\varphi(\lambda)$ 是用 $a_{11}(\lambda)$ 除 $a_{1k}(\lambda)$ 的商，而 $b(\lambda)\neq 0$ 是剩余，它的幂次低于 $a_{11}(\lambda)$ 之幂次．这时，从 $\boldsymbol{A}(\lambda)$ 之第 k 列减去用 $\varphi(\lambda)$ 乘了后的第一列，就得到了一个与 $\boldsymbol{A}(\lambda)$ 等价的 λ-矩阵，其中已用 $b(\lambda)$ 代替了 $a_{1k}(\lambda)$；再将第 1、k 两列交换，我们就得到左上角是 $b(\lambda)$ 而与 $\boldsymbol{A}(\lambda)$ 等价的一个 λ-矩阵，但 $b(\lambda)$ 的幂次已低于 $a_{11}(\lambda)$ 之幂次．

同样，如果不能被 $a_{11}(\lambda)$ 整除的 $a_{ik}(\lambda)$ 出现在 $\boldsymbol{A}(\lambda)$ 之第 1 列，例如是 $a_{k1}(\lambda)$，我们也可以仿上得到同样的结果，这只要将上段中所用的"行""列"两字对调，而将 $a_{1k}(\lambda)$ 改为 $a_{k1}(\lambda)$ 就行了．

至于 $\boldsymbol{A}(\lambda)$ 的第 1 行以及第 1 列的元素都能为 $a_{11}(\lambda)$ 整除，但有某 $a_{ik}(\lambda)$ 不能被 $a_{11}(\lambda)$ 整除的情形，可以用下列方法将这种情形变成上面的情形，因而也就得到同样的结果——事实上，由于这时 $a_{i1}(\lambda)$ 能被 $a_{11}(\lambda)$ 整除，即

$$a_{i1}(\lambda)=\varphi(\lambda)\cdot a_{11}(\lambda),$$

所以第 i 行减去用 $\varphi(\lambda)$ 乘了后的第 1 行，$a_{i1}(\lambda)$ 就变为 0，而元素 $a_{ik}(\lambda)$ 变为 $a'_{ik}(\lambda)=a_{ik}(\lambda)-\varphi(\lambda)\cdot a_{1k}(\lambda)$，它仍然不能被 $a_{11}(\lambda)$ 整除（因 $a_{1k}(\lambda)$ 据题设已能被 $a_{11}(\lambda)$ 整除）．再将这变换后的第 i 行加到第 1 行，由于现在第 i 行第 1 列处的元是 0，所以 $a_{11}(\lambda)$ 不变，而在第 1 行第 k 列处的元是

$$a_{1k}(\lambda)+a'_{ik}(\lambda)=a_{1k}(\lambda)\cdot[1-\varphi(\lambda)]+a_{ik}(\lambda),$$

它显然不能被 $a_{11}(\lambda)$ 整除．这样就归结到上面已经考察过了的情形．

总之，证得了：

引理 如果 λ-矩阵 $\boldsymbol{A}(\lambda)$ 之左上角元素 $a_{11}(\lambda)$ 不恒等于

零，并且 $\boldsymbol{A}(\lambda)$ 中有不能被 $a_{11}(\lambda)$ 整除的元素，那末就有与 $\boldsymbol{A}(\lambda)$ 等价的 λ-矩阵 $\boldsymbol{B}(\lambda)$，其左上角之元素 $b_{11}(\lambda)$ 也不恒等于零且幂次低于 $a_{11}(\lambda)$ 的次数.

根据这一引理，我们可以把 λ-矩阵 $\boldsymbol{A}(\lambda)$ 化成所谓它的标准形：即首先对于秩 $r\neq0$ 的 λ-矩阵 $\boldsymbol{A}(\lambda)$，利用行和列的交换可令左上角元素 $a_{11}(\lambda)\neq0$；如果矩阵中有不被 $a_{11}(\lambda)$ 整除的元素时，由引理可使 $\boldsymbol{A}(\lambda)$ 变为等价的 λ-矩阵，使它左上角的元有较低的幂次且不恒为零；若这个新 λ-矩阵中所有的元还不能都被它左上角元来整除时，我们又可降低这左上角元的幂次．这样做下去，终可得到一个 λ-矩阵 $\boldsymbol{B}(\lambda)$，使其中所有的元都能被左上角元 $b_{11}(\lambda)$ 所整除†.

于是，这时由于 $\boldsymbol{B}(\lambda)$ 之第 1 行中各元都能被 $b_{11}(\lambda)$ 整除，故由第 2、3、…各列分别减去第 1 列乘以适当之多项式，就可使第一行中除 $b_{11}(\lambda)$ 外所有的元都是零．同样，我们也可以使第一列中从第二行开始的各元都是零．因为 $\boldsymbol{B}(\lambda)$ 里各元已都能被 $b_{11}(\lambda)$ 整除，所以在刚才所说的变形后之矩阵中的一切元素当然也都能被 $b_{11}(\lambda)$ 所整除．不失一般性，可令 $b_{11}(\lambda)$ 之首项系数等于 1，象这样形成之多项式用 $I_1(\lambda)$ 表示，于是矩阵 $\boldsymbol{A}(\lambda)$ 就等价于

$$\begin{pmatrix} I_1(\lambda) & 0 & 0 & \cdots & 0 \\ 0 & c_{22}(\lambda) & c_{23}(\lambda) & \cdots & c_{2n}(\lambda) \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & c_{m2}(\lambda) & c_{m3}(\lambda) & \cdots & c_{mn}(\lambda) \end{pmatrix} \tag{1}$$

之形状，这里的一切 $c_{ij}(\lambda)$ 都能被 $I_1(\lambda)$ 整除.

若 $\boldsymbol{A}(\lambda)$ 之秩 $r=1$，显然每个 $c_{ij}(\lambda)\equiv0$；若 $r>1$，则

† 这是因为，当左上角元之幂次至多降低为零，也就是 $b_{11}(\lambda)$ 为常数时，就已可整除 $\boldsymbol{B}(\lambda)$ 的所有元.

$c_{ij}(\lambda)$ 中至少有不恒为零的，于是我们又可以对 $(m-1)\times(n-1)$-矩阵 $(c_{ij}(\lambda))$ 施行类似于曾经在 $\boldsymbol{A}(\lambda)$ 中施行过了的那些运算．但矩阵 $(c_{ij}(\lambda))$ 之初等变换同时又显然可看做是(1)之初等变换，这样，我们又得到一个矩阵，其前二行和前二列的一切元除在主对角线上的外都必等于零，在主对角线上的是 $I_1(\lambda)$ 与 $I_2(\lambda)$，且显然 $I_2(\lambda)$ 还被 $I_1(\lambda)$ 整除，而其他的元都能被 $I_2(\lambda)$ 整除．

如若 $r>2$，我们又可以继续这样做，最后就得到这样一个矩阵，即在对角线上的元顺次为 $I_1(\lambda), I_2(\lambda), \cdots, I_r(\lambda)$，其它位置的元都是零，且每个 $I_i(\lambda)$ 之首项系数都等于 1，$I_{k+1}(\lambda)$ 又能被 $I_k(\lambda)$ 所整除 $(k=1, \cdots, r-1)$．

于是，我们证得了下面的

定理 1（λ-矩阵的标准形）　*凡秩为 r 的 λ-矩阵 $\boldsymbol{A}(\lambda)$ 可用初等变换变成：*

$$\begin{pmatrix} I_1(\lambda) & 0 & 0 & \cdots & 0 & \\ 0 & I_2(\lambda) & 0 & \cdots & 0 & \\ \vdots & \vdots & \vdots & & \vdots & \boldsymbol{O} \\ 0 & 0 & 0 & \cdots & I_r(\lambda) & \\ & & \boldsymbol{O} & & & \boldsymbol{O} \end{pmatrix}, \tag{2}$$

这里，在对角线上的多项式 $I_k(\lambda)$ 之首项系数是 1，并且

$$I_k(\lambda) \mid I_{k+1}(\lambda) \qquad (k=1, \cdots, r-1).$$

因 λ-矩阵 $A(\lambda)$ 与标准形(2)等价，故它们的 i 级子行列式之最高公因式是分别相等的 $(i=1, 2, \cdots, r)$．而从矩阵(2)去求 i 级子行列式的最高公因式，比直接从 $A(\lambda)$ 去求它们要容易得多．在(2)中，因 $I_1(\lambda)$ 能整除所有的元素，故(2)中一切一级子行列式的最高公因式 $D_1(\lambda)$ 等于 $I_1(\lambda)$．又因在

(2)中不恒等于零的只有主子行列式，故(2)中不恒为零的2级子行列式一定是象$I_i(\lambda)\cdot I_j(\lambda)$这样的，因而再根据关于$I_i(\lambda)$间的关系得知$I_1(\lambda)\cdot I_2(\lambda)\mid I_i(\lambda)I_j(\lambda)$，故有$D_2(\lambda)=I_1(\lambda)\cdot I_2(\lambda)$．继续往下推，不难获知

$$\left.\begin{array}{l} D_1(\lambda)\equiv I_1(\lambda), \\ D_2(\lambda)\equiv I_1(\lambda)\cdot I_2(\lambda), \\ D_3(\lambda)\equiv I_1(\lambda)\cdot I_2(\lambda)\cdot I_3(\lambda), \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ D_r(\lambda)\equiv I_1(\lambda)\cdot I_2(\lambda)\cdot I_3(\lambda)\cdots I_r(\lambda). \end{array}\right\} \tag{3}$$

从(3)式还可看出不仅有$D_{i-1}(\lambda)\mid D_i(\lambda)$，而且除得的商等于$I_i(\lambda)$．今后，叫$I_k(\lambda)$为$A(\lambda)$的第$k$个不变因子，为了表示与$\boldsymbol{A}(\lambda)$发生连系，有时又记为$I_k^A(\lambda)$．

定理1解决了引言中的第4个问题，即什么是λ-矩阵的标准形．

注意：λ-矩阵$\boldsymbol{A}(\lambda)$的标准形(2)是唯一的．这是因为，所谓有标准形指的是有满足如定理1中诸条件的多项式$I_k(\lambda)$，因$D_1(\lambda)$，…，$D_r(\lambda)$显然是由λ-矩阵$\boldsymbol{A}(\lambda)$唯一确定的，于是根据(3)式，$I_1(\lambda)$，$I_2(\lambda)$，…，$I_r(\lambda)$亦依次得以唯一确定．

有了λ-矩阵的标准形，就容易解决引言中的第3个问题，即有：

定理2 λ-矩阵$\boldsymbol{A}(\lambda)$与$\boldsymbol{B}(\lambda)$等价的充要条件是：对任何的i，有$D_i^A(\lambda)\equiv D_i^B(\lambda)$．

证明 从$\boldsymbol{A}(\lambda)\simeq\boldsymbol{B}(\lambda)$可得$D_i^A(\lambda)\equiv D_i^B(\lambda)$，$(i=1,\cdots,r)$，这在上面已证过(参见第183页)．故只需证明充分性．首先，将$\boldsymbol{A}(\lambda)$与$\boldsymbol{B}(\lambda)$均化为如定理1中所述的标准形，就得到它们的诸不变因子，分别记为$I_i^A(\lambda)$及$I_i^B(\lambda)$．于是，由(3)

式得

$$D_i^A(\lambda)=I_1^A(\lambda)\cdot I_2^A(\lambda)\cdots I_i^A(\lambda),$$
$$D_i^B(\lambda)=I_1^B(\lambda)\cdot I_2^B(\lambda)\cdots I_i^B(\lambda).$$

然据假设 $D_i^A(\lambda)=D_i^B(\lambda)$, $(i=1, 2, \cdots)$, 从而知 $\boldsymbol{A}(\lambda)$ 与 $\boldsymbol{B}(\lambda)$ 之秩相等(这是因为, 若 $r_A<r_B$, 则 $D_{r_B}^B(\lambda)$ 有意义, 而 $D_{r_B}^A(\lambda)$ 没有意义), 令公共秩为 r, 则它们都各有 r 个不变因子, 由是再顺次地从 $D_1^A(\lambda)=D_1^B(\lambda)$, $\cdots$, $D_r^A(\lambda)=D_r^B(\lambda)$, 可逐步推得 $I_1^A(\lambda)=I_1^B(\lambda)$, $I_2^A(\lambda)=I_2^B(\lambda)$, $\cdots$, $I_r^A(\lambda)=I_r^B(\lambda)$, 即 $\boldsymbol{A}(\lambda)$ 与 $\boldsymbol{B}(\lambda)$ 都和同一个标准形等价, 因之也必有 $\boldsymbol{A}(\lambda)\simeq\boldsymbol{B}(\lambda)$, 定理 2 证完.

定理 2 虽然给出了 $\boldsymbol{A}(\lambda)\simeq\boldsymbol{B}(\lambda)$ 的充要条件, 但这只有理论上的价值: 因求 $D_i(\lambda)$ (其中 $i=1, 2, \cdots$)就必须求出 λ-矩阵中一切不恒为零之 i 级子行列式, 这是计算工作量很大的. 故一般不采用. 一般是采用:

定理 3 *$\boldsymbol{A}(\lambda)\simeq\boldsymbol{B}(\lambda)$ 的充要条件是: 对任意 i, 有 $I_i^A(\lambda)=I_i^B(\lambda)$, 即 $\boldsymbol{A}(\lambda)$ 与 $\boldsymbol{B}(\lambda)$ 有相同的不变因子.*

证明 先证充分性. 从 $I_i^A(\lambda)=I_i^B(\lambda)$, $(i=1, 2, \cdots)$, 即得

$$\boldsymbol{A}(\lambda)\simeq\begin{pmatrix} I_1^A(\lambda) &&&&& \\ & I_2^A(\lambda) &&&& \\ && \ddots &&& \\ &&& I_r^A(\lambda) && \\ &&&& 0 & \\ &&&&& \ddots \\ &&&&&& 0 \end{pmatrix}=\begin{pmatrix} I_1^B(\lambda) &&&&& \\ & I_2^B(\lambda) &&&& \\ && \ddots &&& \\ &&& I_r^B(\lambda) && \\ &&&& 0 & \\ &&&&& \ddots \\ &&&&&& 0 \end{pmatrix}\simeq\boldsymbol{B}(\lambda).$$

再证必要性. 从 $\boldsymbol{A}(\lambda)\simeq\boldsymbol{B}(\lambda)$, 利用定理 2, 可知

$$D_i^A(\lambda)\equiv D_i^B(\lambda),$$

依次令 $i=1,2,\cdots$, 据(3)式, 就得到 $I_i^A(\lambda)\equiv I_i^B(\lambda)$, $(i=1, 2, \cdots)$. 定理 3 证完.

定理 3 避免了烦琐的行列式之计算, 而只利用了初等变换化标准形, 所以实用价值较大.

注意: 定理 2 及 3 的条件中, 都已隐含了秩相等的条件. 特别地, 当 n 级 λ-矩阵 $\boldsymbol{A}(\lambda)$ 为满秩时, 则据初等变换的意义, 知 $\boldsymbol{A}(\lambda)$ 的行列式 $|\boldsymbol{A}(\lambda)|\equiv cI_1^A(\lambda)\cdot I_2^A(\lambda)\cdots I_n^A(\lambda)$ (其中 c 是不等于零的一个常数), 即每个不变因子 $I_i^A(\lambda)$ 为 $|\boldsymbol{A}(\lambda)|$ 之因式; 又因每 $I_i^A(\lambda)$ 是由矩阵 $\boldsymbol{A}(\lambda)$ 唯一确定, 故它们是 $\boldsymbol{A}(\lambda)$ 之不变量, 这也正是叫 $I_i^A(\lambda)$ 为不变因子的由来.

我们知道: 把 $D_i(\lambda)$ 分解出因子 $I_i(\lambda)$ 之步骤, 完全是由所与的 λ-矩阵 $\boldsymbol{A}(\lambda)$ 用初等变换的方法诱导而得, 因之运算的法则可限制为"加、减、乘、除"的四则运算, 故当 $\boldsymbol{A}(\lambda)$ 之元之系数全为有理数(或全为实数)时, 则不变因子 $I_i(\lambda)$ 之系数亦可限制在有理数或实数的范围内. 如果允许利用复数, 则每个不变因子 $I_i(\lambda)$ 一般又得分解为一次因式的幂积, 令为

$$I_k(\lambda)\equiv(\lambda-\lambda_1)^{e_{k1}}(\lambda-\lambda_2)^{e_{k2}}\cdots(\lambda-\lambda_t)^{e_{kt}},$$

$(k=1, 2, \cdots, r)$. 因为

$$I_{k-1}(\lambda)\,|\,I_k(\lambda),$$

所以

$$e_{1j}\leqslant e_{2j}\leqslant\cdots\leqslant e_{rj}, \quad (j=1, 2, \cdots, t).$$

这里应注意的是: 因 $\lambda_1, \lambda_2, \cdots, \lambda_t$ 是 $I_r(\lambda)$ 的一切相异的根, 所以 $e_{r1}, e_{r2}, \cdots, e_{rt}$ 无一为零, 但 $e_{1j}, e_{2j}, \cdots, e_{r-1,j}$ 中 $(j=1, \cdots, t)$ 可能有为 0 的, 而且如果任一 $e_{ij}=0$ 时 $(j=1, 2, \cdots, t;\ i=1, 2, \cdots, r-1)$, 那末也必 $e_{1j}=\cdots=e_{i-1,j}=0$, 例如, 当

$I_1(\lambda)\equiv 1$ 时，就有 $e_{11}=e_{12}=\cdots=e_{1t}=0$. 我们把

$$\begin{array}{llll}(\lambda-\lambda_1)^{e_{11}}, & (\lambda-\lambda_1)^{e_{21}}, & \cdots, & (\lambda-\lambda_1)^{e_{r1}},\\ (\lambda-\lambda_2)^{e_{12}}, & (\lambda-\lambda_2)^{e_{22}}, & \cdots, & (\lambda-\lambda_2)^{e_{r2}},\\ \cdots\cdots & \cdots\cdots & \cdots\cdots & \cdots\cdots \\ (\lambda-\lambda_t)^{e_{1t}}, & (\lambda-\lambda_t)^{e_{2t}}, & \cdots, & (\lambda-\lambda_t)^{e_{rt}},\end{array}$$

中不是常数 1 的因式(即指数 $e_{ik}\neq 0$ 的 $(\lambda-\lambda_k)^{e_{ik}}$) 统统叫做 $A(\lambda)$ 的初等因子．为要表示与 λ_s 有关系起见，又把

$$(\lambda-\lambda_s)^{e_{1s}},\ (\lambda-\lambda_s)^{e_{2s}},\ \cdots,\ (\lambda-\lambda_s)^{e_{rs}},\quad (s=1,\ 2,\ \cdots,\ t)$$

中不为常数 1 的因子称为与 $\lambda-\lambda_s$ 相当的初等因子。

例如，λ-矩阵

$$\boldsymbol{A}(\lambda)=\begin{pmatrix}1 &&&&&\\ & 1 &&&&\\ && (\lambda-2)^5(\lambda-3)^3 &&&\\ &&& (\lambda-2)^5(\lambda-3)^4(\lambda+2) &&\\ &&&& 0 &\\ &&&&& \ddots \\ &&&&&& 0\end{pmatrix}$$

的秩显然等于 4, $I_1(\lambda)\equiv I_2(\lambda)\equiv 1$, $I_3(\lambda)\equiv(\lambda-2)^5(\lambda-3)^3$, $I_4(\lambda)\equiv(\lambda-2)^5(\lambda-3)^4(\lambda+2)$，于是其初等因子有五个：

$$(\lambda-2)^5,\ (\lambda-2)^5,\ (\lambda-3)^3,\ (\lambda-3)^4,\ (\lambda+2).$$

而与 $\lambda-2$ 相当的两个初等因子都是 $(\lambda-2)^5$; 与 $\lambda-3$ 相当的两个初等因子是 $(\lambda-3)^3$ 及 $(\lambda-3)^4$; 与 $\lambda+2$ 相当的一个初等因子为 $\lambda+2$.

当然，若两个λ-矩阵 $\boldsymbol{A}(\lambda)$ 和 $\boldsymbol{B}(\lambda)$ 等价 $(\boldsymbol{A}(\lambda)\simeq\boldsymbol{B}(\lambda))$，根据定理 3, 它们的不变因子相同，因之它们的初等因子也相同，可是逆定理却不一定成立！ 即：两个λ-矩阵虽然有相同的初等因子，它们可能不是等价的，例如：

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & (\lambda-2)(\lambda-3) & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

和

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & (\lambda-2)(\lambda-3) & 0 \end{pmatrix}$$

都是 3×4-λ-矩阵，且它们的初等因子都是 $\lambda-2$ 与 $\lambda-3$，可是前者的秩为 2，后者的秩为 3，可见它们不等价.

但是，当两个 $m\times n$-λ 矩阵 $\boldsymbol{A}(\lambda)$ 与 $\boldsymbol{B}(\lambda)$，除了有相同的初等因子以外，还有相等的秩 r 时，若令例如相当于 $\lambda-\lambda_1$ 的初等因子为

$$(\lambda-\lambda_1)^{e_1},\ (\lambda-\lambda_1)^{e_2},\ \cdots,\ (\lambda-\lambda_1)^{e_k}, \tag{4}$$

这里 $e_1\leqslant e_2\leqslant\cdots\leqslant e_k$，则由初等因子的定义，即知(4)中相当于 $\lambda-\lambda_1$ 的 k 个初等因子必定是 k 个不变因子的因式，所以 $k\leqslant r$. 再因为 $I_{j-1}(\lambda)\,|\,I_j(\lambda)$，$(j=2,\ \cdots,\ r)$，故知这时 $(\lambda-\lambda_1)^{e_k}$ 必为 $I_r^A(\lambda)$ 与 $I_r^B(\lambda)$ 的公因式. 对其它的初等因子，亦可类似地处理. 于是 $I_r^A(\lambda)$ 与 $I_r^B(\lambda)$ 的一次因式的幂积完全一样，不得不有 $I_r^A(\lambda)\equiv I_r^B(\lambda)$. 同样，对于任意的 i，我们有 $I_i^A(\lambda)\equiv I_i^B(\lambda)$. 因之，根据定理 3，即得 $\boldsymbol{A}(\lambda)\simeq\boldsymbol{B}(\lambda)$. 所以，可得到两个 λ-矩阵等价的又一个充要条件如下：

定理 4 $\boldsymbol{A}(\lambda)$ 与 $\boldsymbol{B}(\lambda)$ 为等价的充要条件是：它们的秩相等，并且有相同的初等因子.

关于 λ-矩阵等价的问题，也有类似于前章第 5 节定理 6 之推论的结果，即：

定理 5 λ-矩阵 $\boldsymbol{A}(\lambda)$ 与 $\boldsymbol{B}(\lambda)$ 为等价的充要条件是存在

两个可逆 λ-矩阵 $\boldsymbol{P}(\lambda)$ 与 $\boldsymbol{Q}(\lambda)$，使[†]

$$\boldsymbol{B}(\lambda)=\boldsymbol{P}(\lambda)\cdot\boldsymbol{A}(\lambda)\cdot\boldsymbol{Q}(\lambda).$$

证明 先证必要性．设 $\boldsymbol{A}(\lambda)$ 与 $\boldsymbol{B}(\lambda)$ 是两个 $m\times n$-λ-矩阵．因为 $\boldsymbol{E}_{ij}\cdot\boldsymbol{A}(\lambda)$，$\boldsymbol{E}_i(\alpha)\cdot\boldsymbol{A}(\lambda)$ 与 $\boldsymbol{E}_{ij}(f(\lambda))\cdot\boldsymbol{A}(\lambda)$ 分别是从 $\boldsymbol{A}(\lambda)$ 的行经过一次第 1°、2°、3° 类型的初等变换而得的矩阵，而把 $\boldsymbol{B}_{ij}$，$\boldsymbol{E}_i(\alpha)$，$\boldsymbol{E}_{ij}(f(\lambda))$ 右乘 $\boldsymbol{A}(\lambda)$ 则是对 $\boldsymbol{A}(\lambda)$ 的列初等变换的结果（这里，$\boldsymbol{E}_{ij}$、$\boldsymbol{E}_i(\alpha)$、$\boldsymbol{E}_{ij}(f(\lambda))$ 的意义，参看第 146 页，都叫做初等矩阵，有 $|\boldsymbol{B}_{ij}|=-1$，$|\boldsymbol{E}_i(\alpha)|=\alpha\neq0$，$|\boldsymbol{E}_{ij}(f(\lambda))|=1$，其中 $f(\lambda)$ 为 λ 之多项式，即初等矩阵都是可逆 λ-矩阵）．于是，若 $\boldsymbol{A}(\lambda)\simeq\boldsymbol{B}(\lambda)$，则 $\boldsymbol{B}(\lambda)$ 是从 $\boldsymbol{A}(\lambda)$ 经过有限多次的初等变换而得，这无异乎是说将 $\boldsymbol{A}(\lambda)$ 分别左乘且右乘以若干个（有限的）初等矩阵，故若令左乘与右乘的若干个初等矩阵的积各以 $\boldsymbol{P}(\lambda)$ 与 $\boldsymbol{Q}(\lambda)$ 记之，则得 $\boldsymbol{B}(\lambda)=\boldsymbol{P}(\lambda)\cdot\boldsymbol{A}(\lambda)\cdot\boldsymbol{Q}(\lambda)$，而 $\boldsymbol{P}(\lambda)$ 与 $\boldsymbol{Q}(\lambda)$ 均为有限多个初等矩阵的积，即为可逆 λ-矩阵．

再证充分性．若 $\boldsymbol{B}(\lambda)=\boldsymbol{P}(\lambda)\cdot\boldsymbol{A}(\lambda)\cdot\boldsymbol{Q}(\lambda)$，$\boldsymbol{P}(\lambda)$ 与 $\boldsymbol{Q}(\lambda)$ 都是可逆 λ-矩阵，则因 $|\boldsymbol{P}(\lambda)|=a$（非零之常数），故由 $\boldsymbol{P}(\lambda)$ 之标准形：

$$\boldsymbol{P}(\lambda)\simeq\begin{pmatrix} I_1^P(\lambda) & & \\ & \ddots & \\ & & I_m^P(\lambda)\end{pmatrix},$$

即得 $$a=|\boldsymbol{P}(\lambda)|\equiv kI_1^P(\lambda)\cdot I_2^P(\lambda)\cdots I_m^P(\lambda),$$

这里 k 为非零之常数．于是，比较 λ 之次数，就得

$$I_1^P(\lambda)\equiv\cdots\equiv I_m^P(\lambda)\equiv1,$$

† 若 λ-矩阵 $\boldsymbol{P}(\lambda)$ 为满秩的，但其行列式是一个非零的常数（即为 λ 的零次多项式），那末就叫 $\boldsymbol{P}(\lambda)$ 为可逆 λ-矩阵．由逆矩阵之定义容易得知：$\boldsymbol{P}(\lambda)$ 为可逆 λ-矩阵的充要条件是它的逆矩阵 $\boldsymbol{P}(\lambda)^{-1}$ 也为 λ-矩阵．

这证明了 $\boldsymbol{P}(\lambda)\simeq\boldsymbol{E}$(单位矩阵). 于是, 根据上一段所述, 得知 $\boldsymbol{P}(\lambda)=\boldsymbol{P}_1(\lambda)\cdot\boldsymbol{E}\cdot\boldsymbol{Q}_1(\lambda)$, 但 $\boldsymbol{P}_1(\lambda)$ 与 $\boldsymbol{Q}_1(\lambda)$ 都等于有限多个 $\boldsymbol{E}_{ij}$, $\boldsymbol{E}_i(\alpha)$, $\boldsymbol{E}_{ij}(f(\lambda))$形的矩阵之乘积, 故 $\boldsymbol{P}(\lambda)$也等于有限多个这些矩阵之积. 同理, $\boldsymbol{Q}(\lambda)$ 也如此. 这就是说, $\boldsymbol{B}(\lambda)$ 可以从 $\boldsymbol{A}(\lambda)$ 经过有限多次初等变换而得, 即证明了

$$\boldsymbol{A}(\lambda)\simeq\boldsymbol{B}(\lambda).$$

定理 5 证完.

从上面证明的过程中, 还证得下面二性质:

性质 1 可逆 λ-矩阵与单位矩阵等价.

性质 2 可逆 λ-矩阵等于有限多个初等矩阵的乘积.

定理 2～5 虽然都是判断两个 λ-矩阵等价的准则, 除了定理 2, 定理 5 实际也不用, 原因是求 $\boldsymbol{P}(\lambda)$与 $\boldsymbol{Q}(\lambda)$ 仍是件困难的工作, 固然由 $\boldsymbol{P}(\lambda)$ 与 $\boldsymbol{Q}(\lambda)$ 存在可断定 $\boldsymbol{A}(\lambda)\simeq\boldsymbol{B}(\lambda)$, 但用怎样简便的方法才能知道 $\boldsymbol{P}(\lambda)$与 $\boldsymbol{Q}(\lambda)$存在或否呢? 这是没有解决的. 所以, 有实用价值的是定理 3, 这只要把每个λ-矩阵用有限多次的初等变换化为标准形后再看它们的不变因子是否全同就行了. 有了不变因子, 初等因子的求得也并非总是容易的, 原因是要牵涉到多项式的求根问题, 对某些特殊情形, 能否有较简单方法呢?

所谓某些特殊情形, 指的是 λ-矩阵 $\boldsymbol{A}(\lambda)$ 的形状较简单些. 当然, 分块矩阵, 如

$$\boldsymbol{A}(\lambda)=\begin{pmatrix}\boldsymbol{B}(\lambda) & \boldsymbol{O}\\ \boldsymbol{O} & \boldsymbol{C}(\lambda)\end{pmatrix}$$

(其中 $\boldsymbol{B}(\lambda)$与 $\boldsymbol{C}(\lambda)$都是 λ-矩阵, $\boldsymbol{O}$ 表示零矩阵)是形状较简的一种 λ-矩阵. 于是, 我们希望能从 $\boldsymbol{B}(\lambda)$与 $\boldsymbol{C}(\lambda)$ 的不变因子看出 $\boldsymbol{A}(\lambda)$的不变因子, 因为 $\boldsymbol{B}(\lambda)$与 $\boldsymbol{C}(\lambda)$的行、列之个数都较 $\boldsymbol{A}(\lambda)$ 的少, 求它们的不变因子要容易得多. 可是, 这种

希望很容易被说明是落空的，因为，例如

$$A(\lambda)=\begin{pmatrix}\lambda^2+2\lambda+2 & & & \\ & \lambda^3+\lambda^2-2 & & \\ & & \lambda+1 & \\ & & & (\lambda+1)^2\end{pmatrix},$$

如令

$$B(\lambda)=\begin{pmatrix}\lambda^2+2\lambda+2 & 0\\ 0 & \lambda^3+\lambda^2-2\end{pmatrix},$$

$$C(\lambda)=\begin{pmatrix}\lambda+1 & 0\\ 0 & (\lambda+1)^2\end{pmatrix},$$

很容易看出 $B(\lambda)$ 的不变因子为 $I_1^B(\lambda)=\lambda^2+2\lambda+2$, $I_2^B(\lambda)=\lambda^3+\lambda^2-2$（这是因为 $(\lambda^2+2\lambda+2)(\lambda-1)=\lambda^3+\lambda^2-2$），而 $C(\lambda)$ 的不变因子显为 $I_1^C(\lambda)=\lambda+1$, $I_2^C(\lambda)=(\lambda+1)^2$；但由 $\lambda+1$, $(\lambda+1)^2$, $\lambda^2+2\lambda+2$, $\lambda^3+\lambda^2-2$ 却不能断言 $A(\lambda)$ 的不变因子是什么．这说明了：直接从不变因子的角度考虑问题是有困难的．然从初等因子来考虑，有：

定理 6 设 λ-矩阵

$$A(\lambda)=\begin{pmatrix}B(\lambda) & O\\ O & C(\lambda)\end{pmatrix}$$

为分块形，则 $B(\lambda)$，$C(\lambda)$ 的各个初等因子之全体能构成 $A(\lambda)$ 的全部初等因子．

证明 先将 $B(\lambda)$ 与 $C(\lambda)$ 分别化为标准形：

$$B(\lambda)\simeq\begin{pmatrix}I_1^B(\lambda) & & & & \\ & \ddots & & & \\ & & I_{r_1}^B(\lambda) & & \\ & & & 0 & \\ & & & & \ddots & \\ & & & & & 0\end{pmatrix},$$

$$
C(\lambda) \simeq \begin{pmatrix} I_1^C(\lambda) & & & & & \\ & \ddots & & & & \\ & & I_{r_2}^C(\lambda) & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \end{pmatrix}.
$$

这里，$A(\lambda)$ 的秩 $r=r_1+r_2$. 再令 $I_i^B(\lambda)$, $I_i^C(\lambda)$ 分解为不同的一次因式的幂积，如

$$
I_k^B(\lambda) \equiv (\lambda-\lambda_1)^{e_{k1}}(\lambda-\lambda_2)^{e_{k2}}\cdots(\lambda-\lambda_t)^{e_{kt}},
$$
$$
(k=1,\ 2,\ \cdots,\ r_1),
$$
$$
I_l^C(\lambda) \equiv (\lambda-\lambda_1)^{d_{l1}}(\lambda-\lambda_2)^{d_{l2}}\cdots(\lambda-\lambda_t)^{d_{lt}},
$$
$$
(l=1,\ 2,\ \cdots,\ r_2).
$$

因之，$B(\lambda)$ 和 $C(\lambda)$ 的初等因子分别是

$$
(\lambda-\lambda_1)^{e_{11}},\ \cdots,\ (\lambda-\lambda_1)^{e_{r_1 1}},\ \cdots,\ (\lambda-\lambda_t)^{e_{1t}},\ \cdots,\ (\lambda-\lambda_t)^{e_{r_1 t}}
$$

和

$$
(\lambda-\lambda_1)^{d_{11}},\ \cdots,\ (\lambda-\lambda_1)^{d_{r_2 1}},\ \cdots,\ (\lambda-\lambda_t)^{d_{1t}},\ \cdots,\ (\lambda-\lambda_t)^{d_{r_2 t}}
$$

中不为常数的多项式.

现在来证明 $B(\lambda)$, $C(\lambda)$ 的这些初等因子就是 $A(\lambda)$ 的全部初等因子. 不失普遍性，可以把那些与 $\lambda-\lambda_1$ 相当的 $B(\lambda)$, $C(\lambda)$ 之全部初等因子来考虑. 将

$$
e_{11},\ \cdots,\ e_{r_1 1},\ d_{11},\ \cdots,\ d_{r_2 1}
$$

重新按大小的顺序来排列，令为

$$
c_1,\ c_2,\ \cdots,\ c_r,
$$

即

$$
0 \leqslant c_1 \leqslant c_2 \leqslant \cdots \leqslant c_r.
$$

由 $A(\lambda)$ 的形状容易看出，在矩阵 $B(\lambda)$ 或 $C(\lambda)$ 上施行的初等变换，都可以看做是在 $A(\lambda)$ 上所施行的初等变换. 于是可知

$$
\boldsymbol{A}(\lambda) \simeq \begin{pmatrix} I_1^{B}(\lambda) & & & & & & & \\ & \ddots & & & & & & \\ & & I_{r_1}^{B}(\lambda) & & & & & \\ & & & I_1^{C}(\lambda) & & & & \\ & & & & \ddots & & & \\ & & & & & I_{r_2}^{C}(\lambda) & & \\ & & & & & & 0 & \\ & & & & & & & \ddots \\ & & & & & & & & 0 \end{pmatrix}
$$

$$
\simeq \begin{pmatrix} (\lambda-\lambda_1)^{c_1}\varphi_1(\lambda) & & & & & \\ & (\lambda-\lambda_1)^{c_2}\varphi_2(\lambda) & & & & \\ & & \ddots & & & \\ & & & (\lambda-\lambda_1)^{c_r}\varphi_r(\lambda) & & \\ & & & & 0 & \\ & & & & & \ddots \\ & & & & & & 0 \end{pmatrix},
$$

式中 r 个多项式 $\varphi_1(\lambda), \cdots, \varphi_r(\lambda)$ 都不含因式 $\lambda-\lambda_1$. 因此，含在 $D_r^A(\lambda), D_{r-1}^A(\lambda), \cdots, D_1^A(\lambda)$ 中因式 $\lambda-\lambda_1$ 的最高幂指数分别等于

$$\sum_{i=1}^{r} c_i, \ \sum_{i=1}^{r-1} c_i, \ \cdots, \ \sum_{i=1}^{2} c_i, \ c_1;$$

随而，再据不变因子 $I_k^A(\lambda)$ 的定义，则知含在 $I_r^A(\lambda)$, $I_{r-1}^A(\lambda), \cdots, I_2^A(\lambda), I_1^A(\lambda)$ 中因式 $\lambda-\lambda_1$ 的最高幂指数分别等于 $c_r, c_{r-1}, \cdots, c_2, c_1$. 这就是说：$\boldsymbol{A}(\lambda)$ 中与 $\lambda-\lambda_1$ 相当的初等因子是 $(\lambda-\lambda_1)^{c_1}, (\lambda-\lambda_1)^{c_2}, \cdots, (\lambda-\lambda_1)^{c_r}$ 中不为零幂指数的幂 $(\lambda-\lambda_1)^{c_j}$(即 $C_j \neq 0$)，因而就是 $\boldsymbol{B}(\lambda), \boldsymbol{C}(\lambda)$ 中与 $\lambda-\lambda_1$ 相当的全部初等因子. 同理，对 $\lambda-\lambda_2, \cdots, \lambda-\lambda_t$ 说来，均如此. 于是已证明了 $\boldsymbol{B}(\lambda)$, $\boldsymbol{C}(\lambda)$ 的全部初等因子都是 $\boldsymbol{A}(\lambda)$ 的初等因子.

剩下要证明：除此以外，$\boldsymbol{A}(\lambda)$ 再没有别的初等因子.

首先提请注意的是：上面将每个 $I_i^B(\lambda)$ 与 $I_i^C(\lambda)$ 分解出不同的因式的结果就已经暗示着 $\lambda-\lambda_1, \lambda-\lambda_2, \cdots, \lambda-\lambda_t$ 是包含在 $I_{r_1}^B(\lambda)$ 与 $I_{r_2}^C(\lambda)$ 中的全部不同的根所对应的一次因式，即 $\lambda_1, \cdots, \lambda_t$ 是 $I_{r_1}^B(\lambda) \cdot I_{r_2}^C(\lambda)=0$ 的全部不同的根.

今设 $(\lambda-a)^k$ 是 $\boldsymbol{A}(\lambda)$ 的一个初等因子，于是 $(\lambda-a)^k$ 必为含在某个 $I_i^A(\lambda)$ 中 $\lambda-a$ 的最高次幂，因之 $(\lambda-a)^k \mid I_r^A(\lambda)$, 故 $(\lambda-a)^k \mid D_r^A(\lambda)$, 说明了 $\lambda=a$ 为 $D_r^A(\lambda)$ 之一根，即 $D_r^A(a)=0$. 但另一方面，由

$$\boldsymbol{A}(\lambda) \simeq \begin{pmatrix} I_1^B(\lambda) & & & & & & & \\ & \ddots & & & & & & \\ & & I_{r_1}^B(\lambda) & & & & & \\ & & & I_1^C(\lambda) & & & & \\ & & & & \ddots & & & \\ & & & & & I_{r_2}^C(\lambda) & & \\ & & & & & & 0 & \\ & & & & & & & \ddots \\ & & & & & & & & 0 \end{pmatrix}$$

又可知 $D_r^A(\lambda)=I_1^B(\lambda)\cdots I_{r_1}^B(\lambda)I_1^C(\lambda)\cdots I_{r_2}^C(\lambda)$，故从 $D_r^A(a)=0$ 以及 $I_i^B(\lambda)\mid I_{r_1}^B(\lambda)$ 和 $I_i^C(\lambda)\mid I_{r_2}^C(\lambda)$，即得 $I_{r_1}^B(a)\cdot I_{r_2}^C(a)=0$，故 a 不得不为 $\lambda_1,\cdots,\lambda_t$ 中的某一个，此即表明 $(\lambda-a)^k$ 为与某个 $\lambda-\lambda_i$ 相当的一个初等因子，因之由上段所述，即知 $(\lambda-a)^k$ 必为某 $(\lambda-\lambda_i)^{d_{ji}}$，或某 $(\lambda-\lambda_i)^{e_{ji}}$，$(j=1,\cdots,r_1$ 或 $j=1,\cdots,r_2)$. 即证明了：除了 $\boldsymbol{B}(\lambda)$ 与 $\boldsymbol{C}(\lambda)$ 的全部初等因子外，$\boldsymbol{A}(\lambda)$ 再没有别的初等因子.

定理 6 于是完全获证.

再用归纳法，可以很容易地证明：对一般分块矩阵

$$\boldsymbol{A}(\lambda)=\begin{pmatrix}\boldsymbol{B}_1(\lambda) & & & \\ & \boldsymbol{B}_2(\lambda) & & \\ & & \ddots & \\ & & & \boldsymbol{B}_t(\lambda)\end{pmatrix}, \tag{5}$$

有类似于定理 6 之结论，即

定理 7 若 λ-矩阵 $\boldsymbol{A}(\lambda)$ 可写成(5)形的分块矩阵，则 $\boldsymbol{B}_1(\lambda),\ \boldsymbol{B}_2(\lambda),\ \cdots,\ \boldsymbol{B}_t(\lambda)$ 的各个初等因子的全体，就构成 $\boldsymbol{A}(\lambda)$ 的全部初等因子.

由定理 7 又易推知下面的

推论 假如 λ-矩阵

$$\boldsymbol{A}(\lambda)=\begin{pmatrix}f_1(\lambda) & & & & & \\ & f_2(\lambda) & & & & \\ & & \ddots & & & \\ & & & f_r(\lambda) & & \\ & & & & 0 & \\ & & & & & \ddots \\ & & & & & & 0\end{pmatrix},$$

则 $f_1(\lambda),\ f_2(\lambda),\ \cdots,\ f_r(\lambda)$ 的所有一次因式的幂积构成 $\boldsymbol{A}(\lambda)$ 的全部初等因子.

根据定理 6 或 7，形如(5)的 λ-矩阵之不变因子亦易求得，这只要先求出 $\boldsymbol{B}_1(\lambda)$，$\boldsymbol{B}_2(\lambda)$，…，$\boldsymbol{B}_t(\lambda)$ 的全部初等因子，再决定 $\boldsymbol{A}(\lambda)$ 的秩(它等于一切 $\boldsymbol{B}_i(\lambda)$ 之秩的和)，最后按照如证定理 4 的方法那样，即可间接地决定 $\boldsymbol{A}(\lambda)$ 之不变因子．例如，对于

$$\begin{pmatrix} \lambda-1 & 0 \\ 0 & \lambda \end{pmatrix},$$

根据定理 6，它的初等因子为 λ 与 λ−1，而由其秩等于 2，可推知不变因子为

$$I_1(\lambda)\equiv 1$$

与

$$I_2(\lambda)\equiv\lambda(\lambda-1).$$

对于某些特殊 λ-矩阵，有时直接根据定义也容易求得它的不变因子及初等因子．

［例 1］ 试求 λ-矩阵

$$\boldsymbol{A}(\lambda)=\begin{pmatrix} \lambda-a & c_1 & & & \\ & \lambda-a & c_2 & & \\ & & \ddots & \ddots & \\ & & & \lambda-a & c_{n-1} \\ & & & & \lambda-a \end{pmatrix}$$

的不变因子和初等因子，其中 c_1，c_2，…，c_{n-1} 都是常数，且 $c_1c_2\cdots c_{n-1}\neq 0$.

解 因 $\boldsymbol{A}(\lambda)$ 为 n 级方阵，其行列式显然等于 $(\lambda-a)^n$，即

$$D_n^A(\lambda)=(\lambda-a)^n.$$

但删去 $\boldsymbol{A}(\lambda)$ 之第一列及第 n 行后所剩下之 $(n-1)$ 级子式为非零的常数 $c_1c_2\cdots c_{n-1}$，故

$$D_{n-1}^{A}(\lambda)=1,$$

因之

$$D_1^A(\lambda)\equiv\cdots\equiv D_{n-2}^A(\lambda)\equiv 1,$$

于是 $\boldsymbol{A}(\lambda)$之不变因子为

$$I_1^A(\lambda)\equiv\cdots\equiv I_{n-1}^A(\lambda)\equiv 1,$$

$$I_n^A(\lambda)\equiv(\lambda-a)^n,$$

而初等因子只有一个:

$$(\lambda-a)^n.$$

[例 2] 试求 λ-矩阵

$$\boldsymbol{A}(\lambda)=\begin{pmatrix} \lambda & 0 & 0 & \cdots & 0 & a_n \\ -1 & \lambda & 0 & \cdots & 0 & a_{n-1} \\ 0 & -1 & \lambda & \cdots & 0 & a_{n-2} \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda & a_2 \\ 0 & 0 & 0 & \cdots & -1 & \lambda+a_1 \end{pmatrix}$$

之标准形.

解 将 $\boldsymbol{A}(\lambda)$之第 2、3、…、n 行分别乘以 $\lambda,\ \lambda^2,\ \cdots,\ \lambda^{n-1}$ 后都加到第一行上去,就得到

$$\boldsymbol{A}(\lambda)\simeq\begin{pmatrix} 0 & 0 & 0 & \cdots & 0 & f(\lambda) \\ -1 & \lambda & 0 & \cdots & 0 & a_{n-1} \\ 0 & -1 & \lambda & \cdots & 0 & a_{n-2} \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda & a_2 \\ 0 & 0 & 0 & \cdots & -1 & \lambda+a_1 \end{pmatrix}$$

$$(=\boldsymbol{E}_{21}(\lambda)\cdot\boldsymbol{E}_{31}(\lambda^2)\cdots\boldsymbol{E}_{n1}(\lambda^{n-1})\cdot\boldsymbol{A}(\lambda)),$$

且有行列式

$$|A(\lambda)| = (-1)^{1+n}\cdot f(\lambda)\cdot(-1)^{n-1} = f(\lambda)$$
$$= \lambda^n + a_1\lambda^{n-1} + \cdots + a_{n-2}\lambda^2 + a_{n-1}\lambda + a_n,$$

即

$$D_n^A(\lambda) \equiv f(\lambda).$$

但矩阵 $A(\lambda)$ 之第一行及第 n 列被删去后而剩下的 $n-1$ 级子式为 $(-1)^{n-1}$，故

$$D_{n-1}^A(\lambda) = 1,$$

因之

$$D_1^A(\lambda) = \cdots = D_{n-2}^A(\lambda) = 1,$$

由是有

$$I_1^A(\lambda) = \cdots = I_{n-1}^A(\lambda) = 1$$

及

$$I_n^A(\lambda) = f(\lambda),$$

所以 $A(\lambda)$ 之标准形为

$$\begin{pmatrix} 1 & & & & \\ & 1 & & & \\ & & \ddots & & \\ & & & 1 & \\ & & & & f(\lambda) \end{pmatrix} (\simeq A(\lambda)).$$

练 习 二 十

1. 求 λ-矩阵

$$\begin{pmatrix} \lambda-\lambda_1 & 0 \\ 0 & \lambda-\lambda_2 \end{pmatrix}, \quad \lambda_1 \neq \lambda_2$$

的标准形，不变因子和初等因子.

2. 试求

$$\begin{pmatrix} \lambda^2(\lambda-1)^2 & 0 & 0 & 0 \\ 0 & \lambda(\lambda-1)^3 & 0 & 0 \\ 0 & 0 & \lambda-1 & 0 \\ 0 & 0 & 0 & \lambda \end{pmatrix}$$

的不变因子、初等因子和标准形。

3. 试证：对任何 λ-矩阵 $\boldsymbol{A}(\lambda)$，恒有

$$(D_k^{\boldsymbol{A}}(\lambda))^2 \mid D_{k-1}^{\boldsymbol{A}}(\lambda)\cdot D_{k+1}^{\boldsymbol{A}}(\lambda).$$

4. 证明：

$$\boldsymbol{A}(\lambda)=\left(\begin{array}{ccc|ccc}\lambda-a & 0 & 0 & -1 & 0 & 0\\ 0 & \lambda-a & 0 & 0 & -1 & 0\\ 0 & 0 & \lambda-a & 0 & 0 & -1\\ \hline b^2 & 1 & 0 & \lambda-a & 0 & 0\\ 0 & b^2 & 1 & 0 & \lambda-a & 0\\ 0 & 0 & b^2 & 0 & 0 & \lambda-a\end{array}\right)$$

与

$$\boldsymbol{B}(\lambda)=\left(\begin{array}{ccc|ccc}-1 & 0 & 0 & 0 & 0 & 0\\ 0 & -1 & 0 & 0 & 0 & 0\\ 0 & 0 & -1 & 0 & 0 & 0\\ \hline 0 & 0 & 0 & (\lambda-a)^2+b^2 & 1 & 0\\ 0 & 0 & 0 & 0 & (\lambda-a)^2+b^2 & 1\\ 0 & 0 & 0 & 0 & 0 & (\lambda-a)^2+b^2\end{array}\right)$$

是等价的；再求 $\boldsymbol{A}(\lambda)$ 的初等因子和不变因子，并求标准形(这里，a 与 b 都是实数)。

5. 设 $2n$ 级矩阵

$$\boldsymbol{A}=\begin{pmatrix}\begin{bmatrix}a & -b\\ b & a\end{bmatrix} & 1 & & & \\ & \begin{bmatrix}a & -b\\ b & a\end{bmatrix} & 1 & & \\ & & \begin{bmatrix}a & -b\\ b & a\end{bmatrix} & \ddots & \\ & & & \ddots & 1\\ & & & & \begin{bmatrix}a & -b\\ b & a\end{bmatrix}\end{pmatrix},$$

求 $2n$ 级 λ-矩阵 $\lambda\boldsymbol{E}-\boldsymbol{A}$ 的不变因子，初等因子；并求约旦标准形(约旦标准形的意义，参见第 247 页)。其中 $b\neq0$，a 与 b 均为实数。

第3节 λ-矩阵的多项式理论

这里，主要是想把一个变量的多项式中一些结论推广到λ-矩阵，也就是说，想建立λ-矩阵的多项式之初步理论．为此，下面讨论的矩阵都只限于n级的，即方阵的场合．

1）λ-矩阵的多项式写法

先说一个例子．例如

$$\boldsymbol{A}(\lambda)=\begin{pmatrix}\lambda^3+\lambda+1 & \lambda^2-\lambda+3\\ \lambda-1 & \lambda^3+\lambda^2+2\end{pmatrix}$$

是一个2级λ-矩阵，其元素中λ的最高次数为3，若把$\boldsymbol{A}(\lambda)$之每个元素都形式地写为λ的3次完全式，如

$$\boldsymbol{A}(\lambda)=\begin{pmatrix}\lambda^3+0\lambda^2+\lambda+1 & 0\lambda^3+\lambda^2-\lambda+3\\ 0\lambda^3+0\lambda^2+\lambda-1 & \lambda^3+\lambda^2+0\lambda+2\end{pmatrix},$$

则由矩阵之加法运算，即知

$$\boldsymbol{A}(\lambda)=\begin{pmatrix}1&0\\0&1\end{pmatrix}\lambda^3+\begin{pmatrix}0&1\\0&1\end{pmatrix}\lambda^2+\begin{pmatrix}1&-1\\1&0\end{pmatrix}\lambda+\begin{pmatrix}1&3\\-1&2\end{pmatrix}.$$

或

$$\boldsymbol{A}(\lambda)=\lambda^3\begin{pmatrix}1&0\\0&1\end{pmatrix}+\lambda^2\begin{pmatrix}0&1\\0&1\end{pmatrix}+\lambda\begin{pmatrix}1&-1\\1&0\end{pmatrix}+\begin{pmatrix}1&3\\-1&2\end{pmatrix},$$

这说明$\boldsymbol{A}(\lambda)$可写为λ之多项式的形状，这时，各个系数都是2级的数量矩阵．

一般地，任何n级λ-矩阵$\boldsymbol{M}(\lambda)$总可以写为如

$$M(\lambda)=\lambda^m M_0+\lambda^{m-1}M_1+\cdots+\lambda^2 M_{m-2}+\lambda M_{m-1}+M_m,$$

或 (6)

$$M(\lambda)=M_0\lambda^m+M_1\lambda^{m-1}+\cdots+M_{m-2}\lambda^2+M_{m-1}\lambda+M_m$$

的形状，式中每个系数 $M_i(i=0, 1, \cdots, m)$ 都是 n 级数量矩阵，且 $M_0\neq 0$（零矩阵）．这时，我们叫 $M(\lambda)$ 为 m 次的，而叫(6)为 λ-矩阵 $M(\lambda)$ 的多项式写法．

我们知道：变量 x 的多项式 $f(x)$ 中若以常量 a 去代换 x，所得的值 $f(a)$ 是确定的．于是会想到：若 A 是一个 n 级数量矩阵，问以 A 去代换(6)中的 λ，所得的结果怎样？可否用符号 $M(A)$ 来记呢？

我们知道 λ-矩阵 $M(\lambda)$ 的多项式之写法有两种，可以把 λ 的幂次写在 n 级数量矩阵系数的左方与右方，即(6)的前一种与后一种表示法．从前一种写法看，其结果应为（以 A 代换 λ）

$$A^m M_0+A^{m-1}M_1+\cdots+A^2 M_{m-2}+AM_{m-1}+M_m,$$

而由后一种写法，结果则为

$$M_0A^m+M_1A^{m-1}+\cdots+M_{m-2}A^2+M_{m-1}A+M_m,$$

由于矩阵乘法不满足交换律，故这两种结果一般来讲是不同的，因之不能用符号 $M(A)$ 来记，原因是它的意义不明确．为了使得与一个变量 x 的通常多项式相对应起见，我们用

$$M^{(l)}(A)=A^m M_0+A^{m-1}M_1+\cdots+AM_{m-1}+M_m$$

与

$$M^{(r)}(A)=M_0A^m+M_1A^{m-1}+\cdots+M_{m-1}A+M_m$$

来记，即 λ 的幂次在左方时以 A 代换的结果记为 $M^{(l)}(A)$——l 表示“左”的意思；当 λ 的幂次在右方时以 A 代换的结果记为 $M^{(r)}(A)$——r 表示“右”的意思．

从 λ-矩阵的多项式写法看，容易知道：设两个 λ-矩阵 $M(\lambda)$ 与 $N(\lambda)$ 的次数各为 m 与 k，则 $M(\lambda)\pm N(\lambda)$ 的次数 $\leqslant\max(m, k)$；又因两个非零矩阵之积可能为零矩阵，故

$M(\lambda)\cdot N(\lambda)$之次数$\leqslant m+k$.

2) λ-矩阵中的余式定理

我们知道，一个变量 x 的多项式 $f(x)$，当用 $x-a$ 去除时，所得的商及余式都是唯一的，且余式等于 $f(a)$. 对于这一结论，在λ-矩阵的理论中是否成立呢？

设 $\boldsymbol{A}$ 是一个 n 级数量矩阵，$\boldsymbol{M}(\lambda)$ 是一个 n 级λ-矩阵(m 次的). 所谓"用 $\lambda\boldsymbol{E}-\boldsymbol{A}$ 去除 $\boldsymbol{M}(\lambda)$"这句话，在矩阵的理论中就不很确切. 因为矩阵的乘法不满足交换律，故要谈除法，势必要分清左除与右除. 现在分左除与右除两种情况来讨论.

定理 1 设 $\boldsymbol{M}(\lambda)$ 为一个 m 次的 n 级λ-矩阵，$\boldsymbol{A}$ 是一个 n 级数量矩阵，则必有唯一个 $m-1$ 次的 n 级λ-矩阵 $\boldsymbol{Q}^{(l)}(\lambda)$ 及唯一个 n 级数量矩阵 $\boldsymbol{R}^{(l)}$，满足关系式

$$\boldsymbol{M}(\lambda)=(\lambda\boldsymbol{E}-\boldsymbol{A})\cdot\boldsymbol{Q}^{(l)}(\lambda)+\boldsymbol{R}^{(l)}.$$

同理，也有唯一个 $m-1$ 次的 n 级λ-矩阵 $\boldsymbol{Q}^{(r)}(\lambda)$ 及唯一个 n 级数量矩阵 $\boldsymbol{R}^{(r)}$，满足关系式

$$\boldsymbol{M}(\lambda)=\boldsymbol{Q}^{(r)}(\lambda)\cdot(\lambda\boldsymbol{E}-\boldsymbol{A})+\boldsymbol{R}^{(r)}.$$

并且，实际上还有

$$\boldsymbol{R}^{(l)}=\boldsymbol{M}^{(l)}(\boldsymbol{A}) \quad 与 \quad \boldsymbol{R}^{(r)}=\boldsymbol{M}^{(r)}(\boldsymbol{A}).$$

今后，分别叫 $\boldsymbol{Q}^{(l)}(\lambda)$ 与 $\boldsymbol{R}^{(l)}$ 为以 $\lambda\boldsymbol{E}-\boldsymbol{A}$ 左除 $\boldsymbol{M}(\lambda)$ 的*左商*与*左余*，而分别叫 $\boldsymbol{Q}^{(r)}(\lambda)$ 与 $\boldsymbol{R}^{(r)}$ 为以 $\lambda\boldsymbol{E}-\boldsymbol{A}$ 右除 $\boldsymbol{M}(\lambda)$ 的*右商*与*右余*.

证明 若写 $\boldsymbol{M}(\lambda)=\lambda^m\boldsymbol{M}_0+\lambda^{m-1}\boldsymbol{M}_1+\cdots+\lambda\boldsymbol{M}_{m-1}+\boldsymbol{M}_m$ $(\boldsymbol{M}_0\neq\boldsymbol{0})$，则因

$$\begin{aligned}&\boldsymbol{M}(\lambda)-(\lambda\boldsymbol{E}-\boldsymbol{A})\cdot\lambda^{m-1}\boldsymbol{M}_0\\&\quad=\lambda^{m-1}(\boldsymbol{M}_1+\boldsymbol{A}\boldsymbol{M}_0)+\lambda^{m-2}\boldsymbol{M}_2+\cdots+\lambda\boldsymbol{M}_{m-1}+\boldsymbol{M}_m\end{aligned}$$

至多为 $m-1$ 次的λ-矩阵，故归纳地假定

$$\boldsymbol{M}(\lambda)-(\lambda\boldsymbol{E}-\boldsymbol{A})\cdot\lambda^{m-1}\boldsymbol{M}_0=(\lambda\boldsymbol{E}-\boldsymbol{A})\cdot\boldsymbol{N}(\lambda)+\boldsymbol{R},$$

式中 $\boldsymbol{N}(\lambda)$ 最高为 $m-2$ 次的 λ–矩阵，$\boldsymbol{R}$ 为数量矩阵，则

$$\boldsymbol{M}(\lambda)=(\lambda\boldsymbol{E}-\boldsymbol{A})\cdot(\lambda^{m-1}\boldsymbol{M}_0+\boldsymbol{N}(\lambda))+\boldsymbol{R},$$

而 $\lambda^{m-1}\boldsymbol{M}_0+\boldsymbol{N}(\lambda)$ 显为 $m-1$ 次的 λ–矩阵，即可充当定理中所要求的 $\boldsymbol{Q}^{(l)}(\lambda)$，且 $\boldsymbol{R}$ 可充当 $\boldsymbol{R}^{(l)}$。这就用归纳法证明了 $\boldsymbol{Q}^{(l)}(\lambda)$ 与 $\boldsymbol{R}^{(l)}$ 的存在性。

再设有 $\boldsymbol{M}(\lambda)=(\lambda\boldsymbol{E}-\boldsymbol{A})\cdot\boldsymbol{Q}_1^{(l)}(\lambda)+\boldsymbol{R}_1^{(l)}$，$\boldsymbol{Q}_1^{(l)}(\lambda)$ 为 $m-1$ 次的 λ–矩阵，$\boldsymbol{R}_1^{(l)}$ 为数量矩阵，则易推得

$$(\lambda\boldsymbol{E}-\boldsymbol{A})\cdot[\boldsymbol{Q}_1^{(l)}(\lambda)-\boldsymbol{Q}^{(l)}(\lambda)]=\boldsymbol{R}^{(l)}-\boldsymbol{R}_1^{(l)}. \tag{7}$$

可以证明 $\boldsymbol{Q}_1^{(l)}(\lambda)=\boldsymbol{Q}^{(l)}(\lambda)$（设若不然，则(7)之左边至少是一次的 λ–矩阵，但右边为数量矩阵，这显然是不允许的，这就证得了 $Q_1^{(l)}(\lambda)=Q^{(l)}(\lambda)$），随而也必有 $\boldsymbol{R}_1^{(l)}=\boldsymbol{R}^{(l)}$。即证明了 $\boldsymbol{Q}^{(l)}(\lambda)$ 与 $\boldsymbol{R}^{(l)}$ 的唯一性。

最后，若令

$$\boldsymbol{Q}^{(l)}(\lambda)=\lambda^{m-1}\boldsymbol{N}_0+\lambda^{m-2}\boldsymbol{N}_1+\cdots+\lambda\boldsymbol{N}_{m-2}+\boldsymbol{N}_{m-1},\quad \boldsymbol{N}_0\neq\boldsymbol{0},$$

且每 $\boldsymbol{N}_i$ 为 n 级数量矩阵，则由简单的计算可知

$$\begin{aligned}&(\lambda\boldsymbol{E}-\boldsymbol{A})\cdot\boldsymbol{Q}^{(l)}(\lambda)+\boldsymbol{R}^{(l)}\\&\quad=\lambda^m\boldsymbol{N}_0+\lambda^{m-1}(\boldsymbol{N}_1-\boldsymbol{A}\boldsymbol{N}_0)+\lambda^{m-2}(\boldsymbol{N}_2-\boldsymbol{A}\boldsymbol{N}_1)\\&\qquad+\cdots+\lambda(\boldsymbol{N}_{m-1}-\boldsymbol{A}\boldsymbol{N}_{m-2})+(\boldsymbol{R}^{(l)}-\boldsymbol{A}\boldsymbol{N}_{m-1}),\end{aligned}$$

故再由它等于 $\boldsymbol{M}(\lambda)$ 而比较 λ 之同次幂的系数（n 级数量矩阵），可得到

$$\left.\begin{aligned}&\boldsymbol{M}_0=\boldsymbol{N}_0,\\&\boldsymbol{M}_1=\boldsymbol{N}_1-\boldsymbol{A}\boldsymbol{N}_0,\\&\boldsymbol{M}_2=\boldsymbol{N}_2-\boldsymbol{A}\boldsymbol{N}_1,\\&\cdots\cdots\cdots\cdots\cdots\cdots\\&\boldsymbol{M}_{m-2}=\boldsymbol{N}_{m-2}-\boldsymbol{A}\boldsymbol{N}_{m-3},\\&\boldsymbol{M}_{m-1}=\boldsymbol{N}_{m-1}-\boldsymbol{A}\boldsymbol{N}_{m-2},\\&\boldsymbol{M}_m=\boldsymbol{R}^{(l)}-\boldsymbol{A}\boldsymbol{N}_{m-1}.\end{aligned}\right\} \tag{8}$$

从(8)之第一个式子左乘以 $\boldsymbol{A}^m$，第二个左乘以 $\boldsymbol{A}^{m-1}$，第三个左乘以 $\boldsymbol{A}^{m-2}$，…，第 $m-1$ 个左乘 $\boldsymbol{A}^2$，第 m 个左乘以 $\boldsymbol{A}$，最末式左乘以 $\boldsymbol{E}$，然后统统加起来，就得到

$$\boldsymbol{A}^m\boldsymbol{M}_0+\boldsymbol{A}^{m-1}\boldsymbol{M}_1+\boldsymbol{A}^{m-2}\boldsymbol{M}_2+\cdots+\boldsymbol{A}^2\boldsymbol{M}_{m-2}+\boldsymbol{A}\boldsymbol{M}_{m-1}+\boldsymbol{M}_m$$
$$=\boldsymbol{R}^{(l)},$$

即 $\boldsymbol{R}^{(l)}=\boldsymbol{M}^{(l)}(\boldsymbol{A})$.

同理，可证右商与右余的存在性与唯一性，以及

$$\boldsymbol{R}^{(r)}=\boldsymbol{M}^{(r)}(\boldsymbol{A}).$$

定理 1 证完.

定理 1 称为 λ-矩阵中的余式定理.

再引进另一个符号：设 $f(\lambda)=a_0\lambda^k+a_1\lambda^{k-1}+\cdots+a_{k-1}\lambda+a_k$(每 a_i 为常数)为 λ 的一多项式，则定义

$$f(\boldsymbol{A})=a_0\boldsymbol{A}^k+a_1\boldsymbol{A}^{k-1}+\cdots+a_{k-1}\boldsymbol{A}+a_k\boldsymbol{E},$$

式中 $\boldsymbol{A}$ 是一个 n 级矩阵.

推论 1 设 $f(\lambda)$ 为 λ 的一个 k 次多项式，$\boldsymbol{A}$ 为一个 n 级数量矩阵，则必存在唯一个 $k-1$ 次的 n 级 λ-矩阵 $\boldsymbol{Q}(\lambda)$，使得

$$f(\lambda)\boldsymbol{E}=(\lambda\boldsymbol{E}-\boldsymbol{A})\cdot\boldsymbol{Q}(\lambda)+f(\boldsymbol{A})$$
$$=\boldsymbol{Q}(\lambda)\cdot(\lambda\boldsymbol{E}-\boldsymbol{A})+f(\boldsymbol{A}).$$

证明 当 $f(\lambda)=a_0\lambda^k+a_1\lambda^{k-1}+\cdots+a_{k-1}\lambda+a_k$ 时，就有 $f(\lambda)\cdot\boldsymbol{E}=\lambda^k\cdot a_0\boldsymbol{E}+\lambda^{k-1}\cdot a_1\boldsymbol{E}+\cdots+\lambda\cdot a_{k-1}\boldsymbol{E}+a_k\boldsymbol{E}$. 所以，把 $f(\lambda)\cdot\boldsymbol{E}$ 看做是定理 1 中的 $\boldsymbol{M}(\lambda)$时，应有

$$\boldsymbol{R}^{(l)}=\boldsymbol{M}^{(l)}(\boldsymbol{A})$$
$$=\boldsymbol{A}^k\cdot a_0\boldsymbol{E}+\boldsymbol{A}^{k-1}\cdot a_1\boldsymbol{E}+\cdots+\boldsymbol{A}\cdot a_{k-1}\boldsymbol{E}+a_k\boldsymbol{E},$$

及

$$\boldsymbol{R}^{(r)}=\boldsymbol{M}^{(r)}(\boldsymbol{A})$$
$$=a_0\boldsymbol{E}\cdot\boldsymbol{A}^k+a_1\boldsymbol{E}\cdot\boldsymbol{A}^{k-1}+\cdots+a_{k-1}\boldsymbol{E}\cdot\boldsymbol{A}+a_k\boldsymbol{E},$$

由于纯量矩阵与任何矩阵相乘可以交换，故易知，$M^{(l)}(A)$ 与 $M^{(r)}(A)$ 都为 $a_0A^k+a_1A^{k-1}+\cdots+a_{k-1}A+a_kE=f(A)$.

剩下要解决的是 $Q^{(l)}(\lambda)=Q^{(r)}(\lambda)$. 这是因为，由 $R^{(l)}=R^{(r)}=f(A)$ 即得

$$
\begin{aligned}
f(\lambda)E-f(A)&=(\lambda E-A)\cdot Q^{l}(\lambda)\\
&=Q^{(r)}(\lambda)\cdot(\lambda E-A),
\end{aligned}
$$

而

$$
\begin{aligned}
f(\lambda)\cdot E-f(A)&=a_0E\cdot(\lambda^kE-A^k)\\
&+a_1E\cdot(\lambda^{k-1}E-A^{k-1})+\cdots+a_{k-1}E\cdot(\lambda E-A),
\end{aligned}
$$

且每个

$$
\begin{aligned}
\lambda^tE-A^t&=(\lambda E-A)(\lambda^{t-1}E+\lambda^{t-2}A+\cdots+\lambda A^{t-2}+A^{t-1})\\
&=(\lambda^{t-1}E+\lambda^{t-2}A+\cdots+\lambda A^{t-2}+A^{t-1})(\lambda E-A),
\end{aligned}
$$

因之，以 $\lambda E-A$ 左除或右除 $f(\lambda)E-f(A)$ 的商是一致的，故有 $Q^{(l)}(\lambda)=Q^{(r)}(\lambda)$。证完.

推论 2 *设 $f(\lambda)$ 为 λ 之多项式，A 为一个 n 级数量矩阵，则存在 n 级 λ-矩阵 $Q(\lambda)$，使*

$$(\lambda E-A)\cdot Q(\lambda)=f(\lambda)E$$

或

$$Q(\lambda)\cdot(\lambda E-A)=f(\lambda)\cdot E$$

成立的充要条件是 $f(A)=O$.

这由推论 1 即得.

注意：在推论 2 中，由矩阵之等式 $f(\lambda)R=(\lambda E-A)\cdot Q(\lambda)$ 导出 $f(A)=O$，恐读者误以为可直接将 A 代入矩阵等式中的 λ 就很简单地得到，完全象在初等代数里讲余式定理时那样来处理，用不着象证定理 1 时那样. 这种想法是错误的，因为只从形式上看，以 A 代换 λ 好象从 $\lambda E-A$ 即得 $AE-A=O$（即 $f(A)=(AE-A)\cdot Q(A)=O$），然实际上 $\lambda E-A$

为一个 n 级矩阵，即

$$\lambda\boldsymbol{E}-\boldsymbol{A}=\begin{pmatrix}\lambda-a_{11} & -a_{12} & \cdots & -a_{1n}\\ -a_{21} & \lambda-a_{22} & \cdots & -a_{2n}\\ \cdots & \cdots & \cdots & \cdots\\ -a_{n1} & -a_{n2} & \cdots & \lambda-a_{nn}\end{pmatrix},$$

而以 $\boldsymbol{A}$ 直接去代换 λ，右边变成了

$$\begin{pmatrix}\boldsymbol{A}-a_{11} & -a_{12} & \cdots & -a_{1n}\\ -a_{21} & \boldsymbol{A}-a_{22} & \cdots & -a_{2n}\\ \cdots & \cdots & \cdots & \cdots\\ -a_{n1} & -a_{n2} & \cdots & \boldsymbol{A}-a_{nn}\end{pmatrix},$$

这到底是什么东西呢？无法理解．所以说，从 $f(\lambda)\boldsymbol{E}=(\lambda\boldsymbol{E}-\boldsymbol{A})\cdot\boldsymbol{Q}(\lambda)$ 不能直接以 $\boldsymbol{A}$ 代入推得 $f(\boldsymbol{A})=\boldsymbol{O}$.

3）一次 λ-矩阵的等价关系

这里只讨论形状为 $\lambda\boldsymbol{E}-\boldsymbol{A}$ 的一次 λ-矩阵的等价问题．设 $\boldsymbol{A}$、$\boldsymbol{B}$ 为两个 n 级数量矩阵，如果 $\lambda\boldsymbol{B}-\boldsymbol{A}\simeq\lambda\boldsymbol{E}-\boldsymbol{B}$. 于是，由第 2 节定理 5，知有两个可逆 λ-矩阵 $\boldsymbol{M}(\lambda)$、$\boldsymbol{N}(\lambda)$，使

$$(\lambda\boldsymbol{E}-\boldsymbol{A})=\boldsymbol{M}(\lambda)\cdot(\lambda\boldsymbol{E}-\boldsymbol{B})\cdot\boldsymbol{N}(\lambda),$$

由是 $\boldsymbol{M}(\lambda)^{-1}$ 亦为 λ-矩阵，且上式得改写为

$$\boldsymbol{M}(\lambda)^{-1}\cdot(\lambda\boldsymbol{E}-\boldsymbol{A})=(\lambda\boldsymbol{E}-\boldsymbol{B})\cdot\boldsymbol{N}(\lambda). \tag{9}$$

又据定理 1，知满足

$$\boldsymbol{M}(\lambda)^{-1}=(\lambda\boldsymbol{E}-\boldsymbol{B})\cdot\boldsymbol{Q}(\lambda)+\boldsymbol{E}$$

及

$$\boldsymbol{N}(\lambda)=\boldsymbol{Q}_1(\lambda)\cdot(\lambda\boldsymbol{E}-\boldsymbol{A})+\boldsymbol{R}_1$$

的 λ-矩阵 $\boldsymbol{Q}(\lambda)$ 与 $\boldsymbol{Q}_1(\lambda)$、数量矩阵 $\boldsymbol{R}$ 与 $\boldsymbol{E}_1$ 是存在的．以之代入(9)中，并化简，得

$$\begin{aligned}&(\lambda\boldsymbol{E}-\boldsymbol{B})\cdot[\boldsymbol{Q}(\lambda)-\boldsymbol{Q}_1(\lambda)]\cdot(\lambda\boldsymbol{E}-\boldsymbol{A})\\ &=(\lambda\boldsymbol{E}-\boldsymbol{B})\cdot\boldsymbol{R}_1-\boldsymbol{R}\cdot(\lambda\boldsymbol{E}-\boldsymbol{A})\\ &=\lambda(\boldsymbol{R}_1-\boldsymbol{R})+(\boldsymbol{R}\boldsymbol{A}-\boldsymbol{B}\boldsymbol{R}_1),\end{aligned}$$

可以用反证法证明 $Q(\lambda)=Q_1(\lambda)$，设若不然，则上式左边至少是二次的 λ-矩阵，而右边至多为一次的，这显非所许．故不得不有 $Q(\lambda)=Q_1(\lambda)$．因而

$$\lambda(\boldsymbol{R}_1-\boldsymbol{R})+(\boldsymbol{R}\boldsymbol{A}-\boldsymbol{B}\boldsymbol{R}_1)=\boldsymbol{O},$$

由是又有 $\boldsymbol{R}_1=\boldsymbol{R}$，而得

$$\boldsymbol{R}\boldsymbol{A}=\boldsymbol{R}\boldsymbol{R}. \tag{10}$$

再根据定理 1，可令

$$\boldsymbol{M}(\lambda)=(\lambda\boldsymbol{E}-\boldsymbol{A})\cdot\boldsymbol{Q}_2(\lambda)+\boldsymbol{R}_2,$$

式中 $\boldsymbol{Q}_2(\lambda)$ 为 λ-矩阵，$\boldsymbol{R}_2$ 为数量矩阵，于是有

$$\begin{aligned}\boldsymbol{E}&=\boldsymbol{M}(\lambda)^{-1}\cdot\boldsymbol{M}(\lambda)\\&=[(\lambda\boldsymbol{E}-\boldsymbol{B})\cdot\boldsymbol{Q}(\lambda)+\boldsymbol{R}]\cdot[(\lambda\boldsymbol{E}-\boldsymbol{A})\cdot\boldsymbol{Q}_2(\lambda)+\boldsymbol{R}_2]\\&=(\lambda\boldsymbol{B}-\boldsymbol{B})\cdot[\boldsymbol{Q}(\lambda)\cdot(\lambda\boldsymbol{E}-\boldsymbol{A})\cdot\boldsymbol{Q}_2(\lambda)+\boldsymbol{Q}(\lambda)\cdot\boldsymbol{R}_2]\\&\quad+\boldsymbol{R}\cdot(\lambda\boldsymbol{E}-\boldsymbol{A})\cdot\boldsymbol{Q}_2(\lambda)+\boldsymbol{R}\boldsymbol{R}_2;\end{aligned}$$

而根据(10)式又有

$$\boldsymbol{R}(\lambda\boldsymbol{E}-\boldsymbol{A})=(\lambda\boldsymbol{E}-\boldsymbol{B})\boldsymbol{R},$$

以之代入上式中，就有

$$\begin{aligned}\boldsymbol{E}=(\lambda\boldsymbol{E}-\boldsymbol{B})\cdot[&\boldsymbol{Q}(\lambda)\cdot(\lambda\boldsymbol{E}-\boldsymbol{A})\cdot\boldsymbol{Q}_2(\lambda)\\&+\boldsymbol{Q}(\lambda)\cdot\boldsymbol{R}_2+\boldsymbol{R}\cdot\boldsymbol{Q}_2(\lambda)]+\boldsymbol{R}\boldsymbol{R}_2,\end{aligned}$$

即

$$\boldsymbol{E}-\boldsymbol{R}\boldsymbol{R}_2=(\lambda\boldsymbol{E}-\boldsymbol{B})\cdot\boldsymbol{P}(\lambda), \tag{11}$$

而 $\boldsymbol{P}(\lambda)=\boldsymbol{Q}(\lambda)\cdot(\lambda\boldsymbol{E}-\boldsymbol{A})\cdot\boldsymbol{Q}_2(\lambda)+\boldsymbol{Q}(\lambda)\cdot\boldsymbol{R}_2+\boldsymbol{R}\cdot\boldsymbol{Q}_2(\lambda)$．可以证明 $\boldsymbol{P}(\lambda)=\boldsymbol{O}$(可用反证法，设若不然，则(11)之右边至少为一次的 λ-矩阵，欲与其左边之数量矩阵 $\boldsymbol{E}-\boldsymbol{R}\boldsymbol{R}_2$ 相等是不可能的，因之不得不有 $\boldsymbol{P}(\lambda)=\boldsymbol{O}$)，随而 $\boldsymbol{R}\boldsymbol{R}_2=\boldsymbol{E}$，即 $\boldsymbol{R}$ 是满秩的．故再由(10)式，得 $\boldsymbol{A}=\boldsymbol{R}^{-1}\boldsymbol{B}\boldsymbol{R}$，即 $\boldsymbol{A}\sim\boldsymbol{B}$．这就证明了：由 $\lambda\boldsymbol{E}-\boldsymbol{A}\simeq\lambda\boldsymbol{E}-\boldsymbol{B}$ 可得 $\boldsymbol{A}\sim\boldsymbol{B}$．

反之，从 $\boldsymbol{A}\sim\boldsymbol{B}$，可知存在矩阵 $\boldsymbol{P}$，满足 $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}=\boldsymbol{B}$，故

$P^{-1}\cdot(\lambda E-A)\cdot P=\lambda E-B$. 且 P 之满秩性当然意味着 P 与 P^{-1} 都可视为可逆 λ-矩阵，于是据第 2 节定理 5，就有 $\lambda E-A\simeq\lambda E-B$.

这就证得了：

定理 2 设有两个 n 级数量矩阵 A 与 B，那末 $\lambda E-A\simeq\lambda E-B$ 的充要条件是 $A\sim B$.

练习二十一

1. 设

$$A=\begin{pmatrix}1 & -1\\ 0 & 2\end{pmatrix},\quad B=\begin{pmatrix}4 & 2\\ -3 & -1\end{pmatrix},$$

证明 $\lambda E-A\simeq\lambda E-B$；并求两个可逆 λ-矩阵 $M(\lambda)$ 与 $N(\lambda)$，使

$$\lambda E-A=M(\lambda)\cdot(\lambda E-B)\cdot N(\lambda);$$

然后，利用本节的定理 2 之证明方法，求满秩矩阵 R，使 $A=R^{-1}BR$. (即 $A\sim B$.)

2. 设 $M(\lambda)$ 为 m 次的 λ-矩阵，$B(\lambda)$ 为 k 次的 λ-矩阵，其 λ^k 之系数(数量矩阵)为满秩矩阵，$m>k$，试证：有唯一个 $m-k$ 次的 λ-矩阵 $Q^{(l)}(\lambda)$ 及唯一个次数至多为 $k-1$ 次的 λ-矩阵 $R^{(l)}(\lambda)$，使得

$$M(\lambda)=B(\lambda)\cdot Q^{(l)}(\lambda)+R^{(l)}(\lambda).$$

同样，有唯一个 $m-k$ 次的 λ-矩阵 $Q^{(r)}(\lambda)$ 及唯一个次数至多是 $k-1$ 次的 λ-矩阵 $R^{(r)}(\lambda)$，使得

$$M(\lambda)=Q^{(r)}(\lambda)\cdot B(\lambda)+R^{(r)}(\lambda).$$

第 4 节 特征矩阵

在这章的开头就已经谈过，从这一节起将研究如何利用 λ-矩阵的知识来处理数量矩阵，也就是说对数量矩阵作较深刻的分析与探索.

设 A 是一个 n 级数量矩阵. 欲利用 λ-矩阵的理论来探讨 A 之性质，势必先要把 A 联系到某个与 A 有关联的 λ-矩

阵．为此，作 n 级 λ–矩阵 $\lambda\boldsymbol{E}-\boldsymbol{A}$，叫它为 $\boldsymbol{A}$ 的特征矩阵．$\boldsymbol{A}$ 之特征矩阵 $\lambda\boldsymbol{E}-\boldsymbol{A}$ 的行列式

$$\Delta_A(\lambda)=|\lambda\boldsymbol{E}-\boldsymbol{A}|$$

是 λ 的一个 n 次多项式，叫做 $\boldsymbol{A}$ 的特征多项式，因此不论 $\boldsymbol{A}$ 是满秩或降秩，$\lambda\boldsymbol{E}-\boldsymbol{A}$ 总是满秩 λ–矩阵．

$$\Delta_A(\lambda)=|\lambda\boldsymbol{E}-\boldsymbol{A}|=0$$

叫做 $\boldsymbol{A}$ 的特征方程，它的根叫做 $\boldsymbol{A}$ 的特征值（或者叫做 $\boldsymbol{A}$ 的特征根）．假如 λ_1 是 $\boldsymbol{A}$ 的特征值，那末线性方程组

$$(\lambda_1\boldsymbol{E}-\boldsymbol{A})\begin{pmatrix}x_1\\x_2\\\vdots\\x_n\end{pmatrix}=\boldsymbol{0} \tag{12}$$

的任一非零的解向量 $(x_1, x_2, \cdots, x_n)$ 叫做 $\boldsymbol{A}$ 之特征向量，或更确切些叫做对应于 λ_1 的特征向量，而 λ_1 又叫做对应于特征向量 $(x_1, x_2, \cdots, x_n)$ 的特征值．

注意：$\boldsymbol{A}$ 的特征向量是齐次方程组(12)的解，因之两个矩阵虽有相同的特征值，可有不同的特征向量；反之，虽有相同的特征向量也不一定有相同的特征值．譬如

$$\boldsymbol{A}=\begin{pmatrix}0 & -2\\1 & 3\end{pmatrix},\quad \boldsymbol{B}=\begin{pmatrix}-2 & -2\\6 & 5\end{pmatrix},$$

它们的特征值都是 1 与 2，容易验证 $\boldsymbol{A}$ 对应于 1 与 2 的特征向量分别和 $\boldsymbol{X}_1=(-2, 1)$ 与 $\boldsymbol{X}_2=(-1, +1)$ 成比例；$\boldsymbol{B}$ 对应于 1 与 2 的特征向量分别和 $\boldsymbol{Y}_1=(-2, 3)$ 与 $\boldsymbol{Y}_2=(-1, 2)$ 成比例．显然，$\boldsymbol{Y}_1$ 与 $\boldsymbol{Y}_2$ 都不同 $\boldsymbol{X}_1$ 与 $\boldsymbol{X}_2$ 成比例，所以 $\boldsymbol{A}$，$\boldsymbol{B}$ 没有相同的特征向量．又

$$\boldsymbol{C}=\begin{pmatrix}1 & 0\\0 & 3\end{pmatrix},\quad \boldsymbol{D}=\begin{pmatrix}2 & 0\\0 & 6\end{pmatrix}$$

的特征值各为 1, 3 与 2, 6, 故它们无相同的特征值. 但 $\boldsymbol{C}$ 对应于 1 及 $\boldsymbol{D}$ 对应于 2 的特征向量都是 $(\alpha, 0)$, α 是任何异于零的数; 又 $\boldsymbol{C}$ 对应于 3 及 $\boldsymbol{D}$ 对应于 6 的特征向量都是 $(0,\alpha)$, $\alpha\neq 0$. 所以 $\boldsymbol{C}$ 及 $\boldsymbol{D}$ 有完全相同的特征向量.

要想把 λ-矩阵的知识运用到数量矩阵 $\boldsymbol{A}$ 上, 就是要研究 $\boldsymbol{A}$ 的特征矩阵, 研究下面三个问题:

1. $\boldsymbol{A}$ 与其特征多项式 $\Delta_A(\lambda)=|\lambda\boldsymbol{E}-\boldsymbol{A}|$ 的关系;

2. $\boldsymbol{A}$ 之特征值和与 $\boldsymbol{A}$ 有关的矩阵之特征值间的联系;

3. $\boldsymbol{A}$ 之特征值与相应之特征向量间的种种关系.

首先讨论第一个问题.

1) $\boldsymbol{A}$ 与它的特征多项式 $\Delta_A(\lambda)=|\lambda\boldsymbol{E}-\boldsymbol{A}|$ 的关系.

同普通的某变量之多项式一样, 形如

$$a_0\boldsymbol{A}^m+a_1\boldsymbol{A}^{m-1}+\cdots+a_{m-1}\boldsymbol{A}+a_m\boldsymbol{E}$$

的式子(这里 m 为正整数或零, a_i 是常数)称为矩阵 $\boldsymbol{A}$ 的多项式. 假如

$$f(\lambda)=a_0\lambda^m+a_1\lambda^{m-1}+\cdots+a_{m-1}\lambda+a_m,$$

$\boldsymbol{A}$ 是 n 级矩阵, 我们在前节里已经定义过了

$$f(\boldsymbol{A})=a_0\boldsymbol{A}^m+a_1\boldsymbol{A}^{m-1}+\cdots+a_{m-1}\boldsymbol{A}+a_m\boldsymbol{E},$$

显然这也是 n 级矩阵. 下面的定理是表示矩阵 $\boldsymbol{A}$ 与其特征多项式 $\Delta_A(\lambda)$ 间的一个重要关系.

定理 1(凯莱-哈密尔顿 Cayley-Hamilton) *任意矩阵 $\boldsymbol{A}$ 为其特征多项式的根, 即 $\Delta_A(\boldsymbol{A})=\boldsymbol{O}$(零矩阵).*

证明 $\lambda\boldsymbol{E}-\boldsymbol{A}$ 之伴随矩阵 $(\lambda\boldsymbol{E}-\boldsymbol{A})^*$ 仍是 λ-矩阵, 并且有

$$(\lambda\boldsymbol{E}-\boldsymbol{A})\cdot(\lambda\boldsymbol{E}-\boldsymbol{A})^*=\Delta_A(\lambda)\cdot\boldsymbol{E},$$

故由前节定理 1 之推论 2, 就有 $\Delta_A(\boldsymbol{A})=\boldsymbol{O}$, 证完.

凯莱-哈密尔顿定理表明对于任何 n 级矩阵 $\boldsymbol{A}$, 总可找

着一变量 λ 的多项式 $f(\lambda)$，使 $f(\boldsymbol{A})=\boldsymbol{O}$. 凡是具有性质 $\varphi(\boldsymbol{A})=\boldsymbol{O}$ 的变量 λ 之多项式 $\varphi(\lambda)$，都叫做 $\boldsymbol{A}$ 的零化多项式，又叫 $\varphi(\lambda)$ 使 $\boldsymbol{A}$ 零化. $\boldsymbol{A}$ 之特征多项式 $\Delta_A(\lambda)=|\lambda\boldsymbol{E}-\boldsymbol{A}|$ 当然使 $\boldsymbol{A}$ 零化，但是次数比 $\boldsymbol{A}$ 之级数 n 低的多项式也可能使 $\boldsymbol{A}$ 零化，譬如当 $\boldsymbol{A}=C\boldsymbol{E}$ 为 n 级纯量矩阵时，一次多项式 $\lambda-C$ 就已使 $\boldsymbol{A}$ 零化了. 因此，对任何矩阵 $\boldsymbol{A}$，存在次数最低的零化多项式，若限制其首项系数为 1，则 $\boldsymbol{A}$ 的次数最低的零化多项式是唯一的，叫做 $\boldsymbol{A}$ 的最小多项式，用符号 $m_A(\lambda)$ 来表示.

注意：最小多项式不一定是既约的，因为即使 $m_A(\lambda)=\varphi_1(\lambda)\cdot\varphi_2(\lambda)$，由 $\varphi_1(\boldsymbol{A})\cdot\varphi_2(\boldsymbol{A})=\boldsymbol{O}$，也不能得出 $\varphi_1(\boldsymbol{A})=\boldsymbol{O}$ 或 $\varphi_2(\boldsymbol{A})=\boldsymbol{O}$，因为矩阵在乘法中有零因子.

下面来讨论 $\boldsymbol{A}$ 之最小多项式 $m_A(\lambda)$ 与 $\boldsymbol{A}$ 之任何零化多项式之间的关系.

假定 $f(\lambda)$ 是 $\boldsymbol{A}$ 之零化多项式，$f(\lambda)$ 之次数当然不比 $m_A(\lambda)$ 的低. 以 $m_A(\lambda)$ 除 $f(\lambda)$，得

$$f(\lambda)\equiv q(\lambda)\cdot m_A(\lambda)+r(\lambda),$$

这里，$r(\lambda)\equiv 0$ 或 $r(\lambda)$ 之次数低于 $m_A(\lambda)$ 之次数. 于是，由 $f(\boldsymbol{A})=q(\boldsymbol{A})\cdot m_A(\boldsymbol{A})+r(\boldsymbol{A})$，$f(\boldsymbol{A})=\boldsymbol{O}$ 及 $m_A(\boldsymbol{A})=\boldsymbol{O}$，就得到 $r(\boldsymbol{A})=0$，由于再也没有次数较 $m_A(\lambda)$ 低的多项式可以使 $\boldsymbol{A}$ 零化，就有 $r(\lambda)\equiv 0$，因而 $m_A(\lambda)\,|\,f(\lambda)$.

反之，若 $m_A(\lambda)\,|\,f(\lambda)$，则由 $f(\lambda)\equiv q(\lambda)\cdot m_A(\lambda)$，又得到

$$f(\boldsymbol{A})=q(\boldsymbol{A})\cdot m_A(\boldsymbol{A})=\boldsymbol{O}.$$

于是证得了：

定理 2 多项式 $f(\lambda)$ 是 $\boldsymbol{A}$ 之零化多项式的充要条件为 $m_A(\lambda)\,|\,f(\lambda)$. 特别地，有 $m_A(\lambda)\,|\,\Delta_A(\lambda)$，即 $\boldsymbol{A}$ 之最小多项式为其特征多项式的因式.

既已知 $m_A(\lambda)\mid\Delta_A(\lambda)$，再问商 $\dfrac{\Delta_A(\lambda)}{m_A(\lambda)}$ 是什么呢？

今用 $D_j^A(\lambda)$，$I_j^A(\lambda)$ 分别表示 $\lambda\boldsymbol{E}-\boldsymbol{A}$ 的一切 j 级子式的最高公因式和第 j 个不变因子，于是 $D_j^A(\lambda)$、$I_j^A(\lambda)$ 都是 $\Delta_A(\lambda)$ 的因式（$j\leqslant n$）。我们知道：

$$\Delta_A(\lambda)\equiv D_n^A(\lambda)\equiv D_{n-1}^A(\lambda)\cdot I_n^A(\lambda),$$

$$(\lambda\boldsymbol{E}-\boldsymbol{A})^*=D_{n-1}^A(\lambda)\cdot\boldsymbol{M}(\lambda),$$

其中 $\boldsymbol{M}(\lambda)$ 是 n 级 λ-矩阵，其 n^2 个元素的最高公因式是 1，即它们互质．将上式代入 $(\lambda\boldsymbol{E}-\boldsymbol{A})\cdot(\lambda\boldsymbol{E}-\boldsymbol{A})^*=\Delta_A(\lambda)\cdot\boldsymbol{E}$ 中，即得

$$(\lambda\boldsymbol{E}-\boldsymbol{A})\cdot D_{n-1}^A(\lambda)\cdot\boldsymbol{M}(\lambda)=D_{n-1}^A(\lambda)\cdot I_n^A(\lambda)\cdot\boldsymbol{E},$$

由于 $D_{n-1}^A(\lambda)\neq 0$，就应有

$$(\lambda\boldsymbol{E}-\boldsymbol{A})\cdot\boldsymbol{M}(\lambda)=I_n^A(\lambda)\cdot\boldsymbol{E}. \tag{13}$$

据前节定理 1 之推论 2，可知从(13)式应得

$$I_n^A(\boldsymbol{A})=\boldsymbol{O},$$

故由定理 2，得知

$$m_A(\lambda)\mid I_n^A(\lambda),$$

即有多项式 $q(\lambda)$ 使

$$I_n^A(\lambda)\equiv m_A(\lambda)\cdot q(\lambda). \tag{14}$$

另一方面，据 $m_A(\boldsymbol{A})=\boldsymbol{O}$ 及前节定理 1 之推论 2，又知有 λ-矩阵 $\boldsymbol{L}(\lambda)$，使

$$m_A(\lambda)\cdot\boldsymbol{E}=(\lambda\boldsymbol{E}-\boldsymbol{A})\cdot\boldsymbol{L}(\lambda). \tag{15}$$

故由(13)，(14)，(15)，就知道

$$(\lambda\boldsymbol{E}-\boldsymbol{A})\cdot\boldsymbol{M}(\lambda)=(\lambda\boldsymbol{E}-\boldsymbol{A})\cdot\boldsymbol{L}(\lambda)\cdot q(\lambda),$$

于是，据前节定理 1 中的唯一性，就有

$$\boldsymbol{M}(\lambda)\equiv\boldsymbol{L}(\lambda)\cdot q(\lambda),$$

说明了 $q(\lambda)$ 为 $\boldsymbol{M}(\lambda)$ 中 n^2 个元素之一个公因式，故由于它们

的最高公因式等于 1，则必有

$$q(\lambda)\equiv 1. \tag{16}$$

再合并(14)与(16)，得到

$$I_n^A(\lambda)\equiv m_A(\lambda),$$

随而 $\Delta_A(\lambda)\equiv D_n^A(\lambda)\equiv D_{n-1}^A(\lambda)\cdot I_n^A(\lambda)\equiv D_{n-1}^A(\lambda)\cdot m_A(\lambda)$. 故证得了：

定理 3 n 级矩阵 $\boldsymbol{A}$ 之最小多项式 $m_A(\lambda)$ 等于它的特征矩阵 $\lambda\boldsymbol{E}-\boldsymbol{A}$ 中的第 n 个不变因子 $I_n^A(\lambda)$. 因而最小多项式 $m_A(\lambda)$ 除以特征多项式 $\Delta_A(\lambda)$ 的商等于特征矩阵 $\lambda\boldsymbol{E}-\boldsymbol{A}$ 中所有 $n-1$ 级子式的最高公因式 $D_{n-1}^A(\lambda)$.

由定理 3，易知下面的

推论 1 方阵 $\boldsymbol{A}$ 的特征多项式的根，是它的最小多项式的根.

证明 将 λ-矩阵 $\lambda\boldsymbol{E}-\boldsymbol{A}$ 化为标准形

$$\lambda\boldsymbol{E}-\boldsymbol{A}\simeq\begin{pmatrix} I_1^A(\lambda) & & & \\ & \ddots & & \\ & & I_{n-1}^A(\lambda) & \\ & & & I_n^A(\lambda)\end{pmatrix}$$

时，由第 2 节的定理 5，则知

$$\Delta_A(\lambda)=|\lambda\boldsymbol{E}-\boldsymbol{A}|=c\cdot I_1^A(\lambda)\cdots I_{n-1}^A(\lambda)\cdot I_n^A(\lambda),$$

c 为非零的常数(实际上，这时应有 $c=1$). 若 λ_1 是 $\Delta_A(\lambda)=0$ 之一根，则由 $\Delta_A(\lambda_1)=0$ 知 λ_1 必为某 $I_j^A(\lambda)$ 之根，故再据 $I_j^A(\lambda)\,|\,I_n^A(\lambda)$ 即得 $I_n^A(\lambda_1)=0$，亦即 $m_A(\lambda_1)=0$. 证完.

由推论 1 可知：若 $\Delta_A(\lambda)$ 分解为不同的一次因式之幂积，如

$$\Delta_A(\lambda)\equiv(\lambda-\lambda_1)^{k_1}(\lambda-\lambda_2)^{k_2}\cdots(\lambda-\lambda_t)^{k_t},$$

$k_1+k_2+\cdots+k_t=n$($\boldsymbol{A}$ 之级数)，每 $k_i>0$，且当 $i\neq j$ 时 $\lambda_i\neq\lambda_j$，则

$$m_A(\lambda)=(\lambda-\lambda_1)^{l_1}(\lambda-\lambda_2)^{l_2}\cdots(\lambda-\lambda_t)^{l_t},$$

式中 $1\leqslant l_i\leqslant k_i(i=1,\ 2,\ \cdots,\ t)$. 特别地，当每 $k_i=1$ 时(因之 $t=n$)，也必有每 $l_i=1$，因而 $m_A(\lambda)=\varDelta_A(\lambda)$，即又有：

推论 2 *若矩阵 A 之 n 个特征值互异，则它的最小多项式就等于特征多项式.*

由矩阵 A 与它的特征多项式 $\varDelta_A(\lambda)$ 之关系(主要是凯莱-哈密尔顿定理)引出了 A 的特征多项式的一些性质(定理 2、3 及上推论 1 与 2). 下面再来谈 A 之特征值的有关问题.

2) A 之特征值和与 A 有关的矩阵的特征值间的连系.

设 n 级矩阵 A 的 n 个特征值是 $\lambda_1,\ \lambda_2,\ \cdots,\ \lambda_n$. 先讨论方程 $\varDelta_A(\lambda)=|\lambda E-A|=0$ 的系数间的关系，然后再讨论与 A 有关之矩阵的特征值.

据第一章第 2 节，可将 $\varDelta_A(\lambda)$ 分解为 2^n 个 n 级行列式的和，即

$$\varDelta_A(\lambda)=|\lambda E-A|=\begin{vmatrix}\lambda-a_{11} & 0-a_{12} & \cdots & 0-a_{1n}\\ 0-a_{21} & \lambda-a_{22} & \cdots & 0-a_{2n}\\ \vdots & \vdots & & \vdots\\ 0-a_{n1} & 0-a_{n2} & \cdots & \lambda-a_{nn}\end{vmatrix}$$

$$=|\lambda E|+\sum_{k=1}^{n}\sum_{1\leqslant i_1<\cdots<i_k\leqslant n}D(i_1,\ \cdots,\ i_k),$$

这里 $D(i_1,\ \cdots,\ i_k)$表示 n 级行列式，它的第 i_1、$\cdots$、i_k 列分别是 $-A$ 的第 i_1、$\cdots$、i_k 列，而其余的 $n-k$ 个列(如果存在，即在 $k<n$ 时)分别为纯量矩阵 λE 的相应的列. 但

$$D(i_1,\ \cdots,\ i_k)=(-1)^k\lambda^{n-k}\cdot\varDelta(i_1,\ \cdots,\ i_k),$$

这里 $\varDelta(i_1,\ \cdots,\ i_k)$ 是从 A 中取其第 $i_1,\ \cdots,\ i_k$ 列和行而成的 k 级主子式. 于是,

$$\Delta_A(\lambda)=\lambda^n+\sum_{k=1}^{n}\sum_{1\leq i_1<\cdots<i_k\leq n}(-1)^k\cdot\lambda^{n-k}\cdot\Delta(i_1,\ \cdots,\ i_k)$$

$$=\lambda^n+\sum_{k=1}^{n}(-1)^k\cdot S_k\cdot\lambda^{n-k},$$

式中 S_k 是 $\boldsymbol{A}$ 之一切 k 级主子式的和. 故证得了:

定理 4 $\boldsymbol{A}$ 之特征多项式 $\Delta_A(\lambda)$ 中 $\lambda^r(0\leq r<n)$ 的系数, 等于 $\boldsymbol{A}$ 中一切 $n-r$ 级主子式的和乘以 $(-1)^{n-r}$.

于是 $\Delta_A(\lambda)$ 中 λ^{n-1} 之系数等于 $\boldsymbol{A}$ 之主对角线上元素的和乘以负号, 即为 $-(a_{11}+a_{22}+\cdots+a_{nn})$, 但又因这系数等于 n 个特征值 $\lambda_1,\ \cdots,\ \lambda_n$ 之和的反号, 故有

$$\lambda_1+\lambda_2+\cdots+\lambda_n=a_{11}+a_{22}+\cdots+a_{nn}.$$

$a_{11}+a_{22}+\cdots+a_{nn}$ 通常叫做矩阵 A 的迹, 记以 $\mathrm{tr}\boldsymbol{A}$, 即 $\mathrm{tr}\boldsymbol{A}=\sum_{i=1}^{n}a_{ii}$. 因此, $\mathrm{tr}\boldsymbol{A}=\sum_{i=1}^{n}\lambda_i$.

又, $\Delta_A(\lambda)$ 之常数项一方面等于 $(-1)^n|\boldsymbol{A}|$, 另方面又等于 $(-1)^n\lambda_1\lambda_2\cdots\lambda_n$, 故 $\lambda_1\lambda_2\cdots\lambda_n=|\boldsymbol{A}|$. 因而, 当 $\boldsymbol{A}$ 为满秩时, $\lambda_1,\ \lambda_2,\ \cdots,\ \lambda_n$ 都不为零; $\boldsymbol{A}$ 为降秩时, $\lambda_1,\ \lambda_2,\ \cdots,\ \lambda_n$ 中至少有一为零. 于是, $\boldsymbol{A}$ 为满秩的充要条件是它的特征值无一为零. 由定理 3 的推论 1, 又知 $\boldsymbol{A}$ 为满秩的充要条件是其最小多项式 $m_A(\lambda)$ 没有零为根. 当 $\boldsymbol{A}$ 为降秩时, 若秩等于 $r(<n)$, 则由定理 4 得知 $\Delta_A(\lambda)$ 中 $\lambda^i(i<n-r)$ 的系数都为零, 因此, $\Delta_A(\lambda)$ 中零根之重数至少是 $n-r$. 于是有

推论 1 秩是 r 的 n 级矩阵 $\boldsymbol{A}$ 之特征多项式的零根至少是 $n-r$ 重的.

因 n 级降秩矩阵 $\boldsymbol{A}$ 之伴随矩阵 $\boldsymbol{A}^*$ 的秩至多是 1 (第四章第 5 节定理 2), 故由定理 4 又有:

推论 2 当 n 级矩阵 $\boldsymbol{A}$ 为降秩时, $\boldsymbol{A}^*$ 的 n 个特征值至少有 $n-1$ 个为零, 另一个非零的特征值(如存在)则等于 A_{11}

$+A_{22}+\cdots+A_{nn}$，这里 A_{ii} 是 $\boldsymbol{A}$ 中元素 a_{ii} 的代数余子式.

再来讨论与 $\boldsymbol{A}$ 有关之矩阵(如 $\boldsymbol{A}'$, $\boldsymbol{A}^{-1}$, $\boldsymbol{A}^m$ 及 $k\boldsymbol{A}$)的特征值.

首先来看 $\boldsymbol{A}'$. 由于 $(\lambda\boldsymbol{E}-\boldsymbol{A})'=\lambda\boldsymbol{E}-\boldsymbol{A}'$，即得

$$\Delta_{\boldsymbol{A}}(\lambda)=|\lambda\boldsymbol{E}-\boldsymbol{A}|=|(\lambda\boldsymbol{E}-\boldsymbol{A})'|=|\lambda\boldsymbol{E}-\boldsymbol{A}'|=\Delta_{\boldsymbol{A}'}(\lambda),$$

说明 $\boldsymbol{A}'$ 与 $\boldsymbol{A}$ 有相同的特征多项式，因之也有相同的特征值. 设

$$m_{\boldsymbol{A}}(\lambda)=\lambda^m+c_1\lambda^{m-1}+\cdots+c_{m-1}\lambda+c_m,$$

则

$$\boldsymbol{A}^m+c_1\boldsymbol{A}^{m-1}+\cdots+c_{m-1}\boldsymbol{A}+c_m\boldsymbol{E}=\boldsymbol{O},$$

于是

$$\begin{aligned}\boldsymbol{O}&=(\boldsymbol{A}^m+c_1\boldsymbol{A}^{m-1}+\cdots+c_{m-1}\boldsymbol{A}+c_m\boldsymbol{E})'\\&=\boldsymbol{A}'^m+c_1\boldsymbol{A}'^{m-1}+\cdots+c_{m-1}\boldsymbol{A}'+c_m\boldsymbol{E},\end{aligned}$$

即 $m_{\boldsymbol{A}}(\boldsymbol{A}')=\boldsymbol{O}$，故由定理 2，就有 $m_{\boldsymbol{A}'}(\lambda)\mid m_{\boldsymbol{A}}(\lambda)$. 同理，又有 $m_{\boldsymbol{A}}(\lambda)\mid m_{\boldsymbol{A}'}(\lambda)$. 所以 $m_{\boldsymbol{A}'}(\lambda)=m_{\boldsymbol{A}}(\lambda)$，即证得了：

定理 5 任何矩阵与它的转置矩阵有相同的特征多项式及最小多项式，因而有相同的特征值.

关于 $\boldsymbol{A}^{-1}$, $\boldsymbol{A}^m$ 及 $k\boldsymbol{A}$ 等，先叙述一些预备知识.

如 $\begin{pmatrix}a_{11}&a_{12}&\cdots&a_{1n}\\0&a_{22}&\cdots&a_{2n}\\\vdots&\vdots&\ddots&\vdots\\0&0&\cdots&a_{nn}\end{pmatrix}$ 形的 n 级矩阵(即 $a_{ij}=0$ 当 $i>j$ 时)叫上三角矩阵. 同样，凡 $i<j$ 时 $a_{ij}=0$ 之矩阵 (a_{ij}) 叫做下三角矩阵. 显然，三角(上或下)矩阵之特征值就是它的主对角线上的 n 个元素.

引理 1 上三角矩阵之积仍是上三角的，其主对角线上元素等于各个的对应元之积，因而上三角矩阵 $\boldsymbol{A}$ 之 k 次幂 $\boldsymbol{A}^k$ 为上三角的，其主对角线上的元等于 $\boldsymbol{A}$ 之主对角线上对

应元的 k 次幂.

证明 设 $\boldsymbol{A}=(a_{ij})$ 与 $\boldsymbol{B}=(b_{ij})$ 都是上三角矩阵，即在 $i>j$ 时有 $a_{ij}=0=b_{ij}$，则对 $\boldsymbol{AB}=\boldsymbol{P}=(p_{ij})$，当 $i>j$ 时有

$$p_{ij}=\sum_{t=1}^{n}a_{it}b_{tj}=\sum_{t=1}^{j}a_{it}b_{tj}+\sum_{t=j+1}^{n}a_{it}b_{tj}$$
$$=\sum_{t=1}^{j}0\cdot b_{tj}+\sum_{t=j+1}^{n}a_{it}\cdot 0=0,$$

即 $\boldsymbol{P}$ 也是上三角的. 又，

$$p_{ii}=\sum_{t=1}^{n}a_{it}b_{ti}=\sum_{t=1}^{i-1}a_{it}b_{ti}+\sum_{t=i+1}^{n}a_{it}b_{ti}+a_{ii}b_{ii}=a_{ii}b_{ii}.$$

故引理 1 获证.

同样，关于下三角矩阵，也有类似的结果.

引理 2 对任何 n 级矩阵 $\boldsymbol{A}$，恒可找到一满秩矩阵 $\boldsymbol{P}$，使 $\boldsymbol{P}^{-1}\boldsymbol{AP}$ 为上三角矩阵.

证明 $n=1$ 时显然. 今归纳地假定对 $n-1$ 级矩阵是成立的. 于是，取 $\boldsymbol{A}$ 之一特征值 λ_1，并取 $\boldsymbol{A}$ 之相应于 λ_1 的一特征向量 $\boldsymbol{X}_1$，再作一满秩矩阵 $\boldsymbol{P}_1$ 使 $\boldsymbol{X}_1$ 为第一个列向量，即 $\boldsymbol{P}_1=(\boldsymbol{X}_1,\ \cdots)$［注意：这样的 $\boldsymbol{P}_1$ 多得很］，由矩阵相乘之意义，可知 $\boldsymbol{AP}_1$ 的第一个列向量为 $\boldsymbol{AX}_1=\lambda_1\boldsymbol{X}_1$，因而 $\boldsymbol{P}_1^{-1}\boldsymbol{AP}_1$ 之第一个列向量为

$$\boldsymbol{P}_1^{-1}(\boldsymbol{AX}_1)=\boldsymbol{P}_1^{-1}(\lambda_1\boldsymbol{X}_1)=\lambda_1(\boldsymbol{P}_1^{-1}\boldsymbol{X}_1)=\lambda_1\cdot\begin{pmatrix}1\\0\\\vdots\\0\end{pmatrix},$$

即

$$\boldsymbol{P}_1^{-1}\boldsymbol{AP}_1=\begin{pmatrix}\lambda_1 & *\\ 0 & \\ \vdots & \boldsymbol{A}_1\\ 0 & \end{pmatrix},$$

其中 A_1 为 $n-1$ 级矩阵，$*$ 为具 $n-1$ 个元的行向量，即为 $1\times(n-1)$-矩阵．由归纳法的假定，得知有 $n-1$ 级矩阵 Q_1 使 $Q_1^{-1}A_1Q_1$ 为上三角矩阵，即

$$Q_1^{-1}A_1Q_1=\begin{pmatrix} & * \\ 0 & \end{pmatrix}$$

形．于是，作 n 级矩阵 $P_2=\begin{pmatrix}1 & 0\\ 0 & Q_1\end{pmatrix}$后，则 $P_2^{-1}=\begin{pmatrix}1 & 0\\ 0 & Q_1^{-1}\end{pmatrix}$，且有

$$P_2^{-1}P_1^{-1}AP_1P_2=\begin{pmatrix}1 & 0\\ 0 & Q_1^{-1}\end{pmatrix}\begin{pmatrix}\lambda_1 & *\\ 0 & A_1\end{pmatrix}\begin{pmatrix}1 & 0\\ 0 & Q_1\end{pmatrix}$$

$$=\begin{pmatrix}\lambda_1 & *Q_1\\ 0 & Q_1^{-1}A_1Q_1\end{pmatrix}=\begin{pmatrix}\lambda_1 & *Q_1\\ 0 & 0*\end{pmatrix},$$

显然是上三角矩阵，故取 $P_1P_2=P$ 即可，证完．

若令 $\lambda_1, \lambda_2, \cdots, \lambda_n$ 为 A 之 n 个特征值，则由引理 2 知：存在满秩矩阵 P，使

$$P^{-1}AP=\begin{pmatrix}\lambda_1 & & * \\ & \lambda_2 & \\ & & \ddots & \\ 0 & & & \lambda_n\end{pmatrix},$$

于是，再据引理 1，得知

$$P^{-1}A^kP=(P^{-1}AP)^k=\begin{pmatrix}\lambda_1^k & & *_1 \\ & \lambda_2^k & \\ & & \ddots & \\ 0 & & & \lambda_n^k\end{pmatrix},$$

易证相似矩阵有相同的特征多项式与最小多项式，故 A^k 之特征值为 $\lambda_1^k, \lambda_2^k, \cdots, \lambda_n^k$．同样可证 kA 之特征值为 $k\lambda_1, k\lambda_2, \cdots, k\lambda_n$．一般地，对任一多项式 $f(x)=a_0x^k+a_1x^{k-1}+\cdots+a_{k-1}x+a_k$，因有

$$P^{-1}\cdot f(A)\cdot P=a_0P^{-1}A^kP+a_1P^{-1}A^{k-1}P+\cdots+a_{k-1}P^{-1}AP+a_kP^{-1}EP$$

$$=\begin{pmatrix} f(\lambda_1) & & & * \\ & f(\lambda_2) & & \\ & & \ddots & \\ 0 & & & f(\lambda_n)\end{pmatrix},$$

即$f(A)$之特征值为$f(\lambda_1), f(\lambda_2), \cdots, f(\lambda_n)$，故证得了：

定理 6 设n级矩阵A的n个特征值是$\lambda_1, \lambda_2, \cdots, \lambda_n$，$f(x)$是$x$的任一个多项式，那末$f(A)$的$n$个特征值是$f(\lambda_1), f(\lambda_2), \cdots, f(\lambda_n)$. 因之$kA$的特征值为$k\lambda_1, \cdots, k\lambda_n$；而$A^m$之特征值为$\lambda_1^m, \lambda_2^m, \cdots, \lambda_n^m$（$m$是正整数）.

附注：运用定理6，又可以得到定理3之推论1的另一个证明. 因由定理6，可知$m_A(A)$之特征值是$m_A(\lambda_1), m_A(\lambda_2), \cdots, m_A(\lambda_m)$. 但$m_A(A)=O$，故$m_A(A)$之每一个特征值都是零，即$m_A(\lambda_i)=0(i=1, 2, \cdots, n)$，这就是定理3之推论1.

为了求A^{-1}的特征值，先考虑A^{-1}的构造. 令$m_A(\lambda)=\lambda^m+c_1\lambda^{m-1}+\cdots+c_{m-1}\lambda+c_m$，于是

$$O=m_A(A)=A^m+c_1A^{m-1}+\cdots+c_{m-1}A+c_mE$$
$$=A(A^{m-1}+c_1A^{m-2}+\cdots+c_{m-1}E)+c_mE,$$

因此，若令$\varphi(\lambda)=\lambda^{m-1}+c_1\lambda^{m-2}+\cdots+c_{m-1}$，则有

$$A\cdot\varphi(A)=-c_mE.$$

这样，当A为满秩时，$\varphi(A)=-c_mA^{-1}$，且$c_m\neq 0$（由于$m_A(\lambda)$无零根），故$A^{-1}=-\frac{1}{c_m}\varphi(A)$. 这就证明了：

引理 3 n级满秩矩阵A的逆矩阵A^{-1}，可以写成次数小于n的A的多项式. 即存在次数小于n的多项式$\psi(x)$，使得$A^{-1}=\psi(A)$.

$$\lambda_1^t c_1 \boldsymbol{X}_1+\lambda_2^t c_2 \boldsymbol{X}_2+\cdots+\lambda_k^t c_k \boldsymbol{X}_k=\boldsymbol{0}, \quad (t=1,\ 2,\ \cdots,\ k-1).$$

于是我们把 k 个向量 $c_i\boldsymbol{X}_i(i=1,\ \cdots,\ k)$ 当做未知量而有 k 个线性齐次方程的方程组

$$\left.\begin{array}{l}c_1 \boldsymbol{X}_1+c_2 \boldsymbol{X}_2+\cdots+c_k \boldsymbol{X}_k=\boldsymbol{0}, \\ \lambda_1 c_1 \boldsymbol{X}_1+\lambda_2 c_2 \boldsymbol{X}_2+\cdots+\lambda_k c_k \boldsymbol{X}_k=\boldsymbol{0}, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ \lambda_1^{k-1} c_1 \boldsymbol{X}_1+\lambda_2^{k-1} c_2 \boldsymbol{X}_2+\cdots+\lambda_k^{k-1} c_k \boldsymbol{X}_k=\boldsymbol{0},\end{array}\right\}$$

其系数之行列式

$$D=\left|\begin{array}{cccc}1 & 1 & \cdots & 1 \\ \lambda_1 & \lambda_2 & \cdots & \lambda_k \\ \vdots & \vdots & & \vdots \\ \lambda_1^{k-1} & \lambda_2^{k-1} & \cdots & \lambda_k^{k-1}\end{array}\right|$$

是范得蒙行列式，由于 $i\neq j$ 时有 $\lambda_i\neq\lambda_j$，故 $D\neq 0$，因而引用克莱姆定理的方法就得 $c_i\boldsymbol{X}_i=\boldsymbol{0}$，但 $\boldsymbol{X}_i\neq\boldsymbol{0}$，故必 $c_i=0$，这证得了：

定理8 对应于矩阵 $\boldsymbol{A}$ 的不同的特征值的任一组特征向量是线性无关的.

由定理8(或直接由定义)可知：对应于一矩阵 $\boldsymbol{A}$ 之不同特征值的特征向量是不相同的．但对应于 $\boldsymbol{A}$ 之同一个特征值的特征向量也可能不同，这是因为 $(\lambda_1\boldsymbol{E}-\boldsymbol{A})\boldsymbol{X}=\boldsymbol{0}$ 之非零解向量一般有无限多个．现在研究对于同一个特征值 λ_1 的所有特征向量与 $\boldsymbol{A}$ 的关系.

我们知道：对应于特征值 λ_1 之所有特征向量与零向量都是齐次方程组

$$(\lambda_1 \boldsymbol{E}-\boldsymbol{A})\begin{pmatrix}x_1 \\ \vdots \\ x_n\end{pmatrix}=\begin{pmatrix}0 \\ \vdots \\ 0\end{pmatrix} \tag{18}$$

的解向量，所有这些解向量形成 n 维空间 $\mathfrak{R}_n$ 的一个 k 维子空间 $\mathfrak{S}$，叫做 $\boldsymbol{A}$ 对应于 λ_1 的特征子空间，这时 $\lambda_1\boldsymbol{E}-\boldsymbol{A}$ 的秩是 $n-k$.

现在假定 $\boldsymbol{X}_1^{(1)}, \boldsymbol{X}_2^{(1)}, \cdots, \boldsymbol{X}_k^{(1)}$ 是(18)的一组 k 个线性无关的解向量．当然，$\boldsymbol{A}\boldsymbol{X}_i^{(1)}=\lambda_1\boldsymbol{X}_i^{(1)}$，$(i=1, \cdots, k)$．再选 $n-k$ 个向量 $\boldsymbol{X}_{k+1}, \cdots, \boldsymbol{X}_n$ 使

$$\boldsymbol{X}_1^{(1)}, \boldsymbol{X}_2^{(1)}, \cdots, \boldsymbol{X}_k^{(1)}, \boldsymbol{X}_{k+1}, \cdots, \boldsymbol{X}_n$$

成为 $\mathfrak{R}_n$ 的一组基底(参看第三章)，作 n 级矩阵

$$\boldsymbol{P}=(\boldsymbol{X}_1^{(1)}, \boldsymbol{X}_2^{(1)}, \cdots, \boldsymbol{X}_k^{(2)}, \boldsymbol{X}_{k+1}, \cdots, \boldsymbol{X}_n),$$

于是 $\boldsymbol{P}$ 是满秩的．据矩阵之乘法规则，得到

$$\begin{aligned}\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}=(&\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{X}_1^{(1)}, \boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{X}_2^{(1)}, \cdots, \boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{X}_k^{(1)},\\ &\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{X}_{k+1}, \cdots, \boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{X}_n)\\ =(&\lambda_1\boldsymbol{P}^{-1}\boldsymbol{X}_1^{(1)}, \lambda_1\boldsymbol{P}^{-1}\boldsymbol{X}_2^{(1)}, \cdots, \lambda_1\boldsymbol{P}^{-1}\boldsymbol{X}_k^{(1)},\\ &\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{X}_{k+1}, \cdots, \boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{X}_n);\end{aligned}$$

但 $\boldsymbol{P}^{-1}\boldsymbol{X}_i^{(1)}$ 是单位矩阵 $\boldsymbol{E}$ 的第 i 列，因之 $\lambda_1\boldsymbol{P}^{-1}\boldsymbol{X}_i^{(1)}$ 是纯量矩阵 $\lambda_1\boldsymbol{E}$ 的第 i 列，故代入上式中，就得到

$$\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}=\begin{pmatrix}\begin{matrix}\lambda_1 & 0 & \cdots & 0\\ 0 & \lambda_1 & \cdots & 0\\ \vdots & \vdots & & \vdots\\ 0 & 0 & \cdots & \lambda_1\end{matrix} & * \\ \boldsymbol{O} & \boldsymbol{E}_1\end{pmatrix},$$

式中 $\boldsymbol{B}_1$ 为 $n-k$ 级矩阵，$\boldsymbol{O}$ 是 $(n-k)\times k$-矩阵，$*$ 表示为 $k\times(n-k)$-矩阵．于是，

$$\begin{aligned}\Delta_A(\lambda)&=|\lambda\boldsymbol{E}-\boldsymbol{A}|=|\lambda\boldsymbol{E}-\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}|\\ &=(\lambda-\lambda_1)^k\cdot|\lambda\boldsymbol{E}-\boldsymbol{B}_1|,\end{aligned}$$

这说明 λ_1 至少是 $\boldsymbol{A}$ 的 k 重特征值，因而证得了：

定理 9 矩阵 $\boldsymbol{A}$ 对应于其特征值 λ_1 的特征子空间的维

数不超过 λ_1 在 $\Delta_A(\lambda)$ 中的重数.

定理 9 又常写为：矩阵之任一特征值的几何重数不大于它的代数重数.

特别地，当特征值 λ_1 为单根时(即 λ_1 之代数重数为 1)，则因相应于 λ_1 的特征子空间之维数不小于 1，因而这时就必为 1，故又得到了：

推论 矩阵 $\boldsymbol{A}$ 的特征多项式之单根所对应的特征子空间的维数等于 1.

这里，又产生了两个问题. 第一，既然每个特征值的几何重数≤它的代数重数，于是问等号"="成立的条件怎样？第二，在导出定理 9 之过程中用到了相似矩阵有相同的特征多项式，于是当 $\boldsymbol{A}$, $\boldsymbol{B}$ 中有一为满秩时(例如 $\boldsymbol{A}$ 是满秩)，则 $\boldsymbol{A}^{-1}(\boldsymbol{AB})\boldsymbol{A}=\boldsymbol{BA}$，说明了 $\boldsymbol{AB}$ 与 $\boldsymbol{BA}$ 相似，故 $\boldsymbol{AB}$ 与 $\boldsymbol{BA}$ 有相同的特征多项式；当 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 都是降秩时，$\boldsymbol{AB}$ 与 $\boldsymbol{BA}$ 之特征多项式怎样呢？

我们先解决第二个问题. 第一个问题关系重大，下一节去解决. 今设 $r_A=r(<n)$，则因有两个满秩矩阵 $\boldsymbol{P}$ 和 $\boldsymbol{Q}$，使

$$\boldsymbol{PAQ}=\boldsymbol{E}^{(r)}=\begin{pmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \end{pmatrix} \quad (r\text{ 个 }1)$$

(参看第 82 页定理 3 及第 147 页定理 6 的推论)，故

$$\boldsymbol{PABP}^{-1}=\boldsymbol{PAQ}\cdot\boldsymbol{Q}^{-1}\boldsymbol{BP}^{-1}=\boldsymbol{B}^{(r)}\boldsymbol{C},$$

$$\boldsymbol{C}=\boldsymbol{Q}^{-1}\boldsymbol{BP}^{-1}=(c_{ij}),$$

且

$$\boldsymbol{Q}^{-1}\boldsymbol{BAQ}=\boldsymbol{Q}^{-1}\boldsymbol{BP}^{-1}\cdot\boldsymbol{PAQ}=\boldsymbol{C}\cdot\boldsymbol{E}^{(r)}.$$

但易验证

$$E^{(r)}C=\begin{pmatrix} c_{11} & c_{12} & \cdots & c_{1n} \\ \cdots & \cdots & \cdots & \cdots \\ c_{r1} & c_{r2} & \cdots & c_{rn} \\ & & \boldsymbol{O} & \end{pmatrix},\quad 及\quad CE^{(r)}=\begin{pmatrix} c_{11} & \cdots & c_{1r} & \\ c_{21} & \cdots & c_{2r} & \boldsymbol{O} \\ \vdots & & \vdots & \\ c_{n1} & \cdots & c_{nr} & \end{pmatrix},$$

于是，

$$|\lambda E-E^{(r)}C|=\begin{vmatrix} \lambda-c_{11} & \cdots & -c_{1r} & & \\ \vdots & & \vdots & & * \\ -c_{r1} & \cdots & \lambda-c_{rr} & & \\ & & & \lambda & \\ & 0 & & & \ddots \\ & & & & \lambda \end{vmatrix}$$

$$=\lambda^{n-r}\cdot\begin{vmatrix} \lambda-c_{11} & \cdots & -c_{1r} \\ \vdots & & \\ -c_{r1} & \cdots & \lambda-c_{rr} \end{vmatrix},$$

$$|\lambda E-CE^{(r)}|=\begin{vmatrix} \lambda-c_{11} & \cdots & -c_{1r} & & \\ \vdots & & \vdots & & 0 \\ -c_{r1} & \cdots & \lambda-c_{rr} & & \\ & & & \lambda & \\ & * & & & \ddots \\ & & & & \lambda \end{vmatrix}$$

$$=\lambda^{n-r}\cdot\begin{vmatrix} \lambda-c_{11} & \cdots & -c_{1r} \\ \vdots & & \\ -c_{r1} & \cdots & \lambda-c_{rr} \end{vmatrix},$$

说明了 $\boldsymbol{P}(\boldsymbol{AB})\boldsymbol{P}^{-1}$ 与 $\boldsymbol{Q}^{-1}(\boldsymbol{BA})\boldsymbol{Q}$ 有相同的特征多项式而为 $\lambda^{n-r}\cdot\begin{vmatrix} \lambda-c_{11} & \cdots & -c_{1r} \\ \vdots & & \vdots \\ -c_{r1} & \cdots & \lambda-c_{rr} \end{vmatrix}$，随而分别和它们相似的 $\boldsymbol{AB}$ 与 $\boldsymbol{BA}$ 亦有这个公共的特征多项式，随而也有相同的迹，即 $\mathrm{tr}\boldsymbol{AB}=\mathrm{tr}\boldsymbol{BA}$，故证得了：

定理 10 不论 A, B 为怎样的两个 n 级矩阵，那末 AB 与 BA 恒有相同的特征多项式，因此有相同的特征值与迹，即：$\Delta_{AB}(\lambda)\equiv\Delta_{BA}(\lambda)$，$\mathrm{tr}AB=\mathrm{tr}BA$.

据定理 10，可推广而得：

推论 设 B_1, B_2, $\cdots$, B_k 是 k 个同级矩阵 A_1, A_2, $\cdots$, A_k 的任何循环排列†，那末 $A_1A_2\cdots A_k$ 与 $B_1B_2\cdots B_k$ 有相同的特征多项式，因之有相同的特征值和迹.

附注 读者可能会问：AB 与 BA 恒有相同的最小多项式吗？这只要举一个例子即可明白，答案是否定的. 例如

$$A=\begin{pmatrix}1&0&0\\0&1&0\\0&0&0\end{pmatrix},\quad B=\begin{pmatrix}1&1&1\\1&1&1\\-1&1&0\end{pmatrix},$$

则

$$AB=\begin{pmatrix}1&1&1\\1&1&1\\0&0&0\end{pmatrix}=C,\quad BA=\begin{pmatrix}1&1&0\\1&1&0\\-1&1&0\end{pmatrix}=D,$$

C 与 D 之特征多项式均为 $\lambda^2(\lambda-2)$，故最小多项式或为 $\lambda(\lambda-2)$ 或为 $\lambda^2(\lambda-2)$. 但由计算可知 $C^2=2C$，故 AB 之最小多项式为 $\lambda(\lambda-2)$. 然

$$D^2-2D=\begin{pmatrix}0&0&0\\0&0&0\\2&-2&0\end{pmatrix}\neq O,$$

故 BA 之最小多项式为 $\lambda^2(\lambda-2)$.

[例] 设 $AB=BA$，证明 A 与 B 有公共的特征向量.

证明 取 A 之任一个特征值 λ，并令 A 对应于 λ 的特征

---

† 所谓一组有次序的 m 个字母 x_1, x_2, $\cdots$, x_m 的循环排列，指的是形状为 $x_i, x_{i+1}, \cdots, x_m, x_1, x_2, \cdots, x_{i-1}$ 之排列 $(i=1, 2, \cdots, m)$.

子空间 $\mathfrak{S}_\lambda$ 之维数为 k，而 $X_1, X_2, \cdots, X_k$ 为 $\mathfrak{S}_\lambda$ 的一组基底，于是 $\boldsymbol{AX}_i=\lambda\boldsymbol{X}_i(i=1, 2, \cdots, k)$. 由于 $\boldsymbol{AB}=\boldsymbol{BA}$，可知 $\boldsymbol{ABX}_i=\boldsymbol{BAX}_i=\lambda\boldsymbol{BX}_i$，即 $\boldsymbol{BX}_i\in\mathfrak{S}_\lambda$，故可写为

$$\boldsymbol{BX}_i=c_{1i}\boldsymbol{X}_1+c_{2i}\boldsymbol{X}_2+\cdots+c_{ki}\boldsymbol{X}_k, \quad (i=1, 2, \cdots, k).$$

今令 $\boldsymbol{X}=\sigma_1\boldsymbol{X}_1+\sigma_2\boldsymbol{X}_2+\cdots+\sigma_k\boldsymbol{X}_k\in\mathfrak{S}_\lambda$，故若能选取不全为零的 k 个数 $\sigma_1, \cdots, \sigma_k$，使得 $\boldsymbol{BX}=\mu\boldsymbol{X}$ 成立，则 $\boldsymbol{X}$ 即为我们所要求的 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 的公共特征向量.

但

$$\boldsymbol{BX}=\sum_{i=1}^{k}\sigma_i\boldsymbol{BX}_i=\sum_{i=1}^{k}\sigma_i\sum_{t=1}^{k}c_{ti}\boldsymbol{X}_t=\sum_{t=1}^{k}\left(\sum_{i=1}^{k}c_{ti}\sigma_i\right)\boldsymbol{X}_t,$$

$$\mu\boldsymbol{X}=\sum_{t=1}^{k}\mu\sigma_t\boldsymbol{X}_t,$$

于是欲 $\boldsymbol{BX}=\mu\boldsymbol{X}$，则从 $\boldsymbol{X}_1, \boldsymbol{X}_2, \cdots, \boldsymbol{X}_k$ 的线性无关性，就有

$$\sum_{i=1}^{k}c_{ti}\sigma_i=\mu\sigma_t \quad (t=1, 2, \cdots, k),$$

故若令

$$\boldsymbol{C}=\begin{pmatrix} c_{11} & c_{12} & \cdots & c_{1k} \\ c_{21} & c_{22} & \cdots & c_{2k} \\ \cdots & \cdots & \cdots & \cdots \\ c_{k1} & c_{k2} & \cdots & c_{kk} \end{pmatrix},$$

则上式得简写为如下之矩阵的形式：

$$(\mu\boldsymbol{E}-\boldsymbol{C})\begin{pmatrix}\sigma_1\\ \sigma_2\\ \vdots\\ \sigma_k\end{pmatrix}=\boldsymbol{0}.$$

因 $\boldsymbol{C}$ 为已知，故只需求 $\boldsymbol{C}$ 之任一特征值 μ 后，再决定对应于 μ 的任一个特征向量 $\begin{pmatrix}\sigma_1\\ \sigma_2\\ \vdots\\ \sigma_n\end{pmatrix}$，由之可得 $\boldsymbol{X}$.

练习二十二

1. 假设 $\boldsymbol{A}$ 是 n 级矩阵，$\varphi(\lambda)$ 为非常数的 λ 之多项式，试证：

i) 若 $\varphi(\lambda)\mid m_{\boldsymbol{A}}(\lambda)$，则 $\varphi(\boldsymbol{A})$ 是降秩的；

ii) 若 $\varphi(\lambda)$ 与 $m_{\boldsymbol{A}}(\lambda)$ 之最高公因式为 $d(\lambda)$，则 $\varphi(\boldsymbol{A})$ 与 $d(\boldsymbol{A})$ 有等秩，即 $r_{\varphi(\boldsymbol{A})}=r_{d(\boldsymbol{A})}$；

iii) $\varphi(\boldsymbol{A})$ 为满秩的充要条件是 $\varphi(\lambda)$ 与 $m_{\boldsymbol{A}}(\lambda)$ 互质.

2. 设 $\boldsymbol{A}=\begin{pmatrix}1&0&0\\1&0&1\\0&1&0\end{pmatrix}$，试证关系式 $\boldsymbol{A}^n=\boldsymbol{A}^{n-2}+\boldsymbol{A}^2-\boldsymbol{E}(n\geqslant3)$，并求 $\boldsymbol{A}^{100}$.

3. 设 $\boldsymbol{A}^2=\boldsymbol{E}$，试证 $\boldsymbol{A}$ 之特征值只能为 ±1；若特征值都等于 1，则 $\boldsymbol{A}=\boldsymbol{E}$.

4. 设 $\boldsymbol{A}^k=\boldsymbol{E}$ （k 是正整数），试证 $\lambda\boldsymbol{E}-\boldsymbol{A}$ 的初等因子都是 λ 的一次多项式.

5. 设 $\lambda_1,\lambda_2,\cdots,\lambda_n$ 是 $\boldsymbol{A}=(a_{ij})$ 的 n 个特征值，证明 $\sum\limits_{i=1}^{n}\lambda_i^2=\sum\limits_{i,j=1}^{n}a_{ij}a_{ji}$.

6. 设 $\boldsymbol{X}$ 是 $\boldsymbol{A}$ 的特征向量，证明 $\boldsymbol{X}$ 也是 $f(\boldsymbol{A})$ 的特征向量，这里 $f(x)$ 为 x 之任一多项式. 并问：对应于 $\boldsymbol{X}$ 的 $\boldsymbol{A}$ 之特征值，和 $f(\boldsymbol{A})$ 之特征值有什么关系？

7. 若对任意矩阵 $\boldsymbol{X}$ 都有 $\operatorname{tr}(\boldsymbol{AX})=0$，试证 $\boldsymbol{A}=\boldsymbol{O}$.

8. 设 n 级矩阵 $\boldsymbol{A}$ 之每元素都是 1，试求 $\boldsymbol{A}$ 的特征值，并求 $\boldsymbol{A}$ 之最小多项式.

9. 设 $\boldsymbol{A}=\begin{pmatrix}1&2&3&\cdots&n-1&n\\2&3&4&\cdots&n&n+1\\3&4&5&\cdots&n+1&n+2\\\vdots&\vdots&\vdots&&\vdots&\vdots\\n&n+1&n+2&\cdots&2n-2&2n-1\end{pmatrix}$，求 $\boldsymbol{A}$ 之特征多项式.

10. 设 $\boldsymbol{A}=\begin{pmatrix}2&-1\\1&3\end{pmatrix}$，试计算 $(\boldsymbol{A}^4-5\boldsymbol{A}^3+6\boldsymbol{A}^2+6\boldsymbol{A}-8\boldsymbol{E})^{-1}$.

11. 已知 3 级矩阵 $\boldsymbol{A}$ 的三个特征值为 $1,-1,2$，试将 $\boldsymbol{A}^{2n}$ 表写为 $\boldsymbol{A}$ 的二次式.

12. 设 $\boldsymbol{A}=(a_{ij})$ 之 n 个特征值为 $w_1,w_2,\cdots,w_n$，试求 $\sum w_i^2$.

13. 设 n 级矩阵 $\boldsymbol{A}=(a_{ik})$ 的秩为 r，且 $\boldsymbol{A}^2=\boldsymbol{A}$，试证 $\operatorname{tr}\boldsymbol{A}=\sum\limits_{i=1}^{n}a_{ii}=r$.

第5节　几何重数与代数重数相等和对角矩阵之关系

现在来解决在前节定理9后提出的第一个问题，即当矩阵 $\boldsymbol{A}$ 之每个特征值的几何重数等于它的代数重数时，$\boldsymbol{A}$ 究竟具有怎样的性质呢？谈几何重数，必然牵涉到特征子空间．令 $\lambda_1, \lambda_2, \cdots, \lambda_k$ 是 $\boldsymbol{A}$ 之所有互不相同的特征值，即

$$\Delta_{\boldsymbol{A}}(\lambda)=|\lambda\boldsymbol{E}-\boldsymbol{A}|=(\lambda-\lambda_1)^{m_1}(\lambda-\lambda_2)^{m_2}\cdots(\lambda-\lambda_k)^{m_k},$$

而当 $i\neq j$ 时 $\lambda_i\neq\lambda_j$，于是 $m_1+m_2+\cdots+m_k=n$；并令 $\mathfrak{S}_{\lambda_i}$ 为 $\boldsymbol{A}$ 关于 λ_i 的特征子空间 $(i=1, \cdots, k)$，$\mathfrak{S}_{\lambda_i}$ 的维数设为 r_i，即 $\lambda_i\boldsymbol{E}-\boldsymbol{A}$ 的秩 $=n-r_i$．

引理1　$\mathfrak{S}_{\lambda_1}\cup\mathfrak{S}_{\lambda_2}\cup\cdots\cup\mathfrak{S}_{\lambda_k}=\mathfrak{S}$ 之维数为 $r_1+r_2+\cdots+r_k$．

证明　若 $\boldsymbol{X}_1+\boldsymbol{X}_2+\cdots+\boldsymbol{X}_k=\boldsymbol{0}(\boldsymbol{X}_i\in\mathfrak{S}_{\lambda_i})$，则逐次地用 $\boldsymbol{E}, \boldsymbol{A}, \boldsymbol{A}^2, \cdots, \boldsymbol{A}^{k-1}$ 去左乘，并利用 $\boldsymbol{A}\boldsymbol{X}_i=\lambda_i\boldsymbol{X}_i$，就得到

$$\left.\begin{array}{l}\boldsymbol{X}_1+\boldsymbol{X}_2+\cdots+\boldsymbol{X}_k=\boldsymbol{0},\\ \lambda_1\boldsymbol{X}_1+\lambda_2\boldsymbol{X}_2+\cdots+\lambda_k\boldsymbol{X}_k=\boldsymbol{0},\\ \lambda_1^2\boldsymbol{X}_1+\lambda_2^2\boldsymbol{X}_2+\cdots+\lambda_k^2\boldsymbol{X}_k=\boldsymbol{0},\\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\ \lambda_1^{k-1}\boldsymbol{X}_1+\lambda_2^{k-1}\boldsymbol{X}_2+\cdots+\lambda_k^{k-1}\boldsymbol{X}_k=\boldsymbol{0},\end{array}\right\}$$

因这时有

$$\begin{vmatrix}1 & 1 & \cdots & 1\\ \lambda_1 & \lambda_2 & \cdots & \lambda_k\\ \cdots & \cdots & \cdots & \cdots\\ \lambda_1^{k-1} & \lambda_2^{k-1} & \cdots & \lambda_k^{k-1}\end{vmatrix}\neq 0,$$

故与证前节定理8的方法一样，可推得 $\boldsymbol{X}_1=\boldsymbol{X}_2=\cdots=\boldsymbol{X}_k=\boldsymbol{0}$．于是，$\mathfrak{S}$ 的每个向量 $\boldsymbol{X}$ 写为

$$X=X_1+X_2+\cdots+X_k \quad (X_i\in\mathfrak{S}_{\lambda_i})$$
的方法只有一种，因之 $\mathfrak{S}_{\lambda_1}$, $\mathfrak{S}_{\lambda_2}$, $\cdots$, $\mathfrak{S}_{\lambda_k}$ 中各个 $\mathfrak{S}_{\lambda_i}$ 的基底之合并即为 $\mathfrak{S}$ 之一组基底，这就说明 $\mathfrak{S}$ 的维数 $=r_1+r_2+\cdots+r_k$，证完.

有了这引理 1，就容易解决每个特征值之几何重数等于代数重数的充要条件.

(一)先设每 $r_i=m_i(i=1, 2, \cdots, k)$，即 A 之每特征值的几何重数等于它的代数重数.

这时，若令 m_i 维空间 $\mathfrak{S}_{\lambda_i}$ 的一组基底为 $X_1^{(i)}$, $X_2^{(i)}$, $\cdots$, $X_{m_i}^{(i)}(i=1, 2, \cdots, k)$. 以它们为列向量，作矩阵
$$P=(X_1^{(1)}, \cdots, X_{m_1}^{(1)}; X_1^{(2)}, \cdots, X_{m_2}^{(2)}; \cdots; X_1^{(k)}, \cdots, X_{m_k}^{(k)}),$$
则 P 有 n 个列，即 P 为 n 级矩阵，且由引理 1 知 P 是满秩的. 如证前节定理 9 的方法一样，可知
$$\begin{aligned}P^{-1}AP&=P^{-1}\cdot(AX_1^{(1)}, \cdots, AX_{m_1}^{(1)}; \cdots; AX_1^{(k)}, \cdots, AX_{m_k}^{(k)})\\&=P^{-1}\cdot(\lambda_1X_1^{(1)}, \cdots, \lambda_1X_{m_1}^{(1)}; \cdots; \lambda_kX_1^{(k)}, \cdots, \lambda_kX_{m_k}^{(k)})\\&=(\lambda_1P^{-1}X_1^{(1)}, \cdots, \lambda_1P^{-1}X_{m_1}^{(1)}; \cdots;\\&\qquad\lambda_kP^{-1}X_1^{(k)}, \cdots, \lambda_kP^{-1}X_{m_k}^{(k)})\\&=\begin{pmatrix}\overbrace{\lambda_1 \cdots \lambda_1}^{m_1\text{个}} & & \\ & \ddots & \\ & & \overbrace{\lambda_k \cdots \lambda_k}^{m_k\text{个}}\end{pmatrix},\end{aligned}$$
证明了 A 与对角矩阵相似.

(二)再设 A 已与对角矩阵相似，，如 $A\sim\begin{pmatrix}\mu_1 & & 0\\ & \mu_2 & \\ & & \ddots & \\ 0 & & & \mu_n\end{pmatrix}$.

于是 μ_1, $\cdots$, μ_n 中有 m_1 个为 λ_1, $\cdots$, m_k 个为 λ_k.

对角矩阵之对角线上元素 $\mu_1, \cdots, \mu_n$ 的位置不论怎样调换，结果仍与原矩阵相似．这是因为，例如

$$E_{12}^{-1}\begin{pmatrix}\mu_1 & & & \\ & \mu_2 & & \\ & & \ddots & \\ & & & \mu_n\end{pmatrix}E_{12}=E_{12}^{-1}\begin{pmatrix}0 & \mu_1 & 0 & \cdots & 0\\ \mu_2 & 0 & 0 & \cdots & 0\\ 0 & 0 & \mu_3 & \cdots & 0\\ \vdots & \vdots & \vdots & & \vdots\\ 0 & 0 & 0 & \cdots & \mu_n\end{pmatrix}$$

$$=\begin{pmatrix}\mu_2 & 0 & 0 & \cdots & 0\\ 0 & \mu_1 & 0 & \cdots & 0\\ 0 & 0 & \mu_3 & \cdots & 0\\ \vdots & \vdots & \vdots & & \vdots\\ 0 & 0 & 0 & \cdots & \mu_n\end{pmatrix}$$

$$=\begin{pmatrix}\mu_2 & & & & 0\\ & \mu_1 & & & \\ & & \mu_3 & & \\ & & & \ddots & \\ 0 & & & & \mu_n\end{pmatrix}.$$

于是，不失一般性，可令

$$A\sim\begin{pmatrix}\overbrace{\lambda_1 \cdots \lambda_1}^{m_1\text{个}} & & & \\ & \ddots & & \\ & & \overbrace{\lambda_k \cdots \lambda_k}^{m_k\text{个}}\end{pmatrix},$$

即有满秩矩阵 $P=(X_1X_2\cdots X_n)$——P 之第 i 列记为 X_i，使[†]

[†] 对角矩阵 $D=\begin{pmatrix}a_1 & & 0\\ & a_2 & \\ & & \ddots \\ 0 & & & a_n\end{pmatrix}$ 简写为 $D=\mathrm{diag}\{a_1, a_2, \cdots, a_n\}$.

$$P^{-1}AP=\begin{pmatrix} \overbrace{\lambda_1 \cdot \cdot \cdot \lambda_1}^{m_1\text{个}} & & \\ & \cdot \cdot \cdot & \\ & & \overbrace{\lambda_k \cdot \cdot \cdot \lambda_k}^{m_k\text{个}} \end{pmatrix}$$

$$=\mathrm{diag}\{\overbrace{\lambda_1, \cdots, \lambda_1}^{m_1\text{个}}, \cdots, \overbrace{\lambda_k, \cdots, \lambda_k}^{m_k\text{个}}\},$$

故 $\boldsymbol{AP}=\boldsymbol{P}\cdot\mathrm{diag}\{\overbrace{\lambda_1, \cdots, \lambda_1}^{m_1\text{个}}, \cdots, \overbrace{\lambda_k, \cdots, \lambda_k}^{m_k\text{个}}\}$.

由矩阵之乘法的意义，得知 $\boldsymbol{AP}$ 的前 m_1 个列向量顺次为 $\boldsymbol{AX}_1, \cdots, \boldsymbol{AX}_{m_1}$，而 $\boldsymbol{P}\cdot\mathrm{diag}\{\underbrace{\lambda_1, \cdots, \lambda_1}_{m_1\text{个}}, \cdots, \underbrace{\lambda_k, \cdots, \lambda_k}_{m_k\text{个}}\}$之前 m_1 个列向量顺次为 $\lambda_1\boldsymbol{X}_1, \cdots, \lambda_1\boldsymbol{X}_{m_1}$，于是就有

$$\boldsymbol{AX}_i=\lambda_1\boldsymbol{X}_i \quad (i=1, 2, \cdots, m_1),$$

说明了 $\boldsymbol{X}_1, \cdots, \boldsymbol{X}_{m_1}$ 都是对应于特征值 λ_1 的 $\boldsymbol{A}$ 之特征向量，故由于它们线性无关(因是满秩矩阵 $\boldsymbol{P}$ 之 m_1 个列)即知特征子空间 $\mathfrak{S}_{\lambda_1}$ 的维数 $r_1 \geqslant m_1$. 但又因 $r_1 \leqslant m_1$(前节的定理 9)，故有 $r_1=m_1$. 同理可证 $r_i=m_i(i=1, 2, \cdots, k)$. 这就是说：$\boldsymbol{A}$ 之每个特征值的几何重数等于它的代数重数.

于是证得了：

定理 1 矩阵 $\boldsymbol{A}$ 之每个特征值的几何重数等于它的代数重数的充要条件是 $\boldsymbol{A}$ 与对角矩阵相似.

定理 1 牵涉到与对角矩阵相似. 但相似矩阵之产生与仿射空间中线性变换的矩阵因基底改变而改变有关(参看第四章第 6 节). 仿射空间中，基底改变的目的在于把线性变换的矩阵尽量简化，以便于研究. 较简单的矩阵当然要算对角矩阵，于是，给了一个矩阵，能否把它化为与它相似的对角矩阵

呢？当然，这并非一定可能．假如可能，它的充要条件是什么？定理 1 也就回答了这个问题，即充要条件是每个特征值的几何重数等于它的代数重数．下面，对这个问题作进一步的探讨．

当每个 $r_i=m_i(i=1, 2, \cdots, k)$ 时，则由引理 1 知 $\sum_{i=1}^{k} r_i=\sum_{i=1}^{k} m_i=n$，即 $\mathfrak{S}=\mathfrak{S}_{\lambda_1}\cup\mathfrak{S}_{\lambda_2}\cup\cdots\cup\mathfrak{S}_{\lambda_k}$ 的维数为 n，亦即矩阵 $\boldsymbol{A}$ 有 n 个线性无关的特征向量．反之，当 $\mathfrak{S}$ 之维数 $=n$，即 $\sum_{i=1}^{k} r_i=n$ 时，则因每 $r_i\leqslant m_i$ 且 $\sum_{i=1}^{k} m_i=n$，故这时必有每 $r_i=m_i$．于是又得到

推论 1 *n 级矩阵 $\boldsymbol{A}$ 与对角矩阵相似的充要条件是 $\boldsymbol{A}$ 有 n 个线性无关的特征向量．*

又，据定理 1 之证明方法，还有：

推论 2 *当 $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}=\boldsymbol{D}$ 是对角矩阵时，$\boldsymbol{P}$ 的 n 个列向量就是 $\boldsymbol{A}$ 的 n 个线性无关的特征向量，它们顺次对应的特征值就是 $\boldsymbol{D}$ 之对角线上的元素．*

特别地，当每 $m_i=1$ 时，即 $\boldsymbol{A}$ 之特征值全是单根时，则由前节定理 9 的推论知每 $r_i=1$，因之每 $r_i=m_i$，故又有

推论 3 *若 $\boldsymbol{A}$ 之特征值全为单根时，则 $\boldsymbol{A}$ 与对角矩阵相似．*

推论 3 的逆显然不成立．就是说，$\boldsymbol{A}$ 虽然与对角矩阵相似，但 $\boldsymbol{A}$ 之特征值可能是重根，这只要看一下对角线上的元素相同的对角矩阵就明白了．这也就是说：$\boldsymbol{A}$ 之特征多项式无重根时，$\boldsymbol{A}$ 一定与对角矩阵相似；但是当 $\boldsymbol{A}$ 与对角矩阵相似时，$\boldsymbol{A}$ 之特征多项式不一定没有重根．而对最小多项式又怎样呢？

因若 $A\sim D=\mathrm{diag}\{\lambda_1, \lambda_2, \cdots, \lambda_n\}$，则有满秩矩阵 P 使 $P^{-1}AP=D$，于是易证：不论 $f(x)$ 为 x 的任何多项式，由于 $cD^k=c(P^{-1}AP)^k=P^{-1}(cA^k)P$，$c$ 为常量，可知

$$P^{-1}\cdot f(A)\cdot P=f(D)=\mathrm{diag}\{f(\lambda_1), \cdots, f(\lambda_n)\}.$$

故若令 $f(x)=(x-\lambda_1)(x-\lambda_2)\cdots(x-\lambda_k)$，但 $\lambda_1, \lambda_2, \cdots, \lambda_k$ 是 A 之所有不同的特征值，则这时有 $f(\lambda_i)=0(i=1, 2, \cdots, k)$，因而 $P^{-1}\cdot f(A)\cdot P=0$，$f(A)=0$，$m_A(\lambda)\mid f(\lambda)$。但 A 之任一特征值又必为 $m_A(\lambda)$ 之根，随而 $m_A(\lambda_i)=0(i=1, 2, \cdots, k)$，故不得不有 $m_A(\lambda)=f(\lambda)$，即 $m_A(\lambda)$ 无重根. 这就证明了：A 与对角矩阵相似时，$m_A(\lambda)$ 必无重根.

反之，若 $m_A(\lambda)=(\lambda-\lambda_1)(\lambda-\lambda_2)\cdots(\lambda-\lambda_k)$ 无重根，则因 $0=m_A(A)=(A-\lambda_1E)(A-\lambda_2E)\cdots(A-\lambda_kE)$，故反复引用前章末的例 2 的结果就得到：

$$\begin{aligned}0=r_{m_A(A)}&\geqslant r_{(A-\lambda_1E)}+r_{(A-\lambda_2E)\cdots(A-\lambda_kE)}-n\geqslant\\ &\geqslant r_{(A-\lambda_1E)}+r_{(A-\lambda_2E)}+r_{(A-\lambda_3E)\cdots(A-\lambda_kE)}-2n\geqslant\\ &\geqslant\cdots\cdots\geqslant\sum_{i=1}^{k}r_{(A-\lambda_iE)}-(k-1)n.\end{aligned}$$

因之，

$$\sum_{i=1}^{k}r_{(A-\lambda_iE)}\leqslant(k-1)n.$$

但因 $r_{(A-\lambda_iE)}=n-r_i$，代入上式中，就得到

$$\sum_{i=1}^{k}r_{(A-\lambda_iE)}=kn-\sum_{i=1}^{k}r_i\leqslant(k-1)n,$$

于是有

$$\sum_{i=1}^{k}r_i\geqslant n=\sum_{i=1}^{k}m_i.$$

再与每 $r_i\leqslant m_i$ 合并，就得到每 $r_i=m_i$. 因而，据定理 1，知 A 能与对角矩阵相似.

又，前节定理 3 说明 $m_A(\lambda)$ 为 $\lambda E-A$ 的最后一个不变

因了 $I_n^A(\lambda)$，故 $m_A(\lambda)$ 无重根(即 $I_n^A(\lambda)$ 无重根)，随而 $\lambda\boldsymbol{E}-\boldsymbol{A}$ 之初等因子都是 λ 的一次式．故证得了：

定理 2 矩阵 $\boldsymbol{A}$ 与对角矩阵相似的充要条件是它的最小多项式 $m_A(\lambda)$ 没有重根，或它的特征矩阵 $\lambda\boldsymbol{E}-\boldsymbol{A}$ 的初等因子都是 λ 的一次式．

定理 2 说明：要判断 $\boldsymbol{A}$ 能否与对角阵相似，就要看它的最小多项式有无重根，但特征多项式的根是最小多项式的根，因此判断 $m_A(\lambda)$ 有无重根，就是要判断特征值在什么时候是 $m_A(\lambda)$ 的单根．

今设特征值 λ_1 是 $m_A(\lambda)$ 的一个单根，即 $(\lambda-\lambda_1)\mid m_A(\lambda)$，$(\lambda-\lambda_1)^2\nmid m_A(\lambda)$，所以 $(\lambda-\lambda_1)^2$ 与 $m_A(\lambda)$ 之最高公因式等于 $(\lambda-\lambda_1)$，因此存在两个多项式 $p(\lambda)$ 和 $q(\lambda)$，使得

$$\lambda-\lambda_1\equiv p(\lambda)\cdot(\lambda-\lambda_1)^2+q(\lambda)\cdot m_A(\lambda),$$

于是 $\boldsymbol{A}-\lambda_1\boldsymbol{E}=p(\boldsymbol{A})\cdot(\boldsymbol{A}-\lambda_1\boldsymbol{E})^2$，从而有 $r_{(A-\lambda_1E)}\leqslant r_{(A-\lambda_1E)^2}$．再与明显的结果 $r_{(A-\lambda_1E)^2}\leqslant r_{(A-\lambda_1E)}$ 比较，就得

$$r_{(A-\lambda_1E)^2}=r_{(A-\lambda_1E)}.$$

反过来，设 λ_1 为 $\boldsymbol{A}$ 之一特征值，并有 $r_{(A-\lambda_1E)^2}=r_{(A-\lambda_1E)}$．于是，显然有 $(\lambda-\lambda_1)\mid m_A(\lambda)$．下面用反证法来证明 λ_1 为 $m_A(\lambda)$ 的单根，设若不然，即有 $(\lambda-\lambda_1)^2\mid m_A(\lambda)$，则 $m_A(\lambda)\equiv g(\lambda)\cdot(\lambda-\lambda_1)^2$，因而由于 $g(\lambda)\cdot(\lambda-\lambda_1)$ 之次数低于 $m_A(\lambda)$ 之次数，则知 $g(\boldsymbol{A})\cdot(\boldsymbol{A}-\lambda_1\boldsymbol{E})\neq\boldsymbol{0}$，故有不为零的列向量 $\boldsymbol{Y}$ 使得 $g(\boldsymbol{A})\cdot(\boldsymbol{A}-\lambda_1\boldsymbol{E})\cdot\boldsymbol{Y}\neq\boldsymbol{0}$[†]，即 $(\boldsymbol{A}-\lambda_1\boldsymbol{E})\cdot g(\boldsymbol{A})\cdot\boldsymbol{Y}\neq\boldsymbol{0}$．但

$$(\boldsymbol{A}-\lambda_1\boldsymbol{E})^2\cdot g(\boldsymbol{A})\cdot\boldsymbol{Y}=m_A(\boldsymbol{A})\cdot\boldsymbol{Y}=\boldsymbol{0}.$$

† 如这 $\boldsymbol{Y}$ 不存在，则对单位矩阵 $\boldsymbol{E}=(\boldsymbol{e}_1, \boldsymbol{e}_2, \cdots, \boldsymbol{e}_n)$ 之每个列向量 $\boldsymbol{e}_i$ 都有 $(\boldsymbol{A}-\lambda_1\boldsymbol{E})g(\boldsymbol{A})\cdot\boldsymbol{e}_i=\boldsymbol{0}$，因之 $(\boldsymbol{A}-\lambda_1\boldsymbol{E})\cdot g(\boldsymbol{A})\cdot\boldsymbol{E}=\boldsymbol{0}$，所以有 $(\boldsymbol{A}-\lambda_1\boldsymbol{E})\cdot g(\boldsymbol{A})=\boldsymbol{0}$．

这是说，线性方程组

$$(\boldsymbol{A}-\lambda_1\boldsymbol{E})^2\cdot\begin{pmatrix}x_1\\x_2\\\vdots\\x_n\end{pmatrix}=\boldsymbol{0} \tag{19}$$

的解 $\boldsymbol{X}=g(\boldsymbol{A})\cdot\boldsymbol{Y}$ 不是方程组

$$(\boldsymbol{A}-\lambda_1\boldsymbol{E})\cdot\begin{pmatrix}x_1\\x_2\\\vdots\\x_n\end{pmatrix}=\boldsymbol{0} \tag{20}$$

之解. 可是，另一方面，从 $r_{(A-\lambda_1E)^2}=r_{(A-\lambda_1E)}$，易知(19)之解空间的维数等于(20)之解空间的维数；且(20)之解又显为(19)之解，即(20)之解空间是(19)之解空间的子空间. 这样，这两个解空间一致，亦即(19)之解亦必为(20)之解，这又与上述矛盾了，由此表明假设 $(\lambda-\lambda_1)^2|m_A(\lambda)$ 是不成立的，即有 $(\lambda-\lambda_1)^2\nmid m_A(\lambda)$，这就证得了 λ_1 是 $m_A(\lambda)$ 之单根. 于是就得到了：

引理2 *$\boldsymbol{A}$ 之特征值 λ_1 是最小多项式 $m_A(\lambda)$ 的单根的充要条件为 $\boldsymbol{A}-\lambda_1\boldsymbol{E}$ 与 $(\boldsymbol{A}-\lambda_1\boldsymbol{E})^2$ 有相同的秩.*

据此，由定理2又得到了：

定理3 *$\boldsymbol{A}$ 与对角矩阵相似的充要条件是：对于 $\boldsymbol{A}$ 的每一个特征值 λ_i，两个降秩矩阵 $\boldsymbol{A}-\lambda_i\boldsymbol{E}$ 和 $(\boldsymbol{A}-\lambda_i\boldsymbol{E})^2$ 有相等的秩，或方程组(19)与(20)为同解方程组.*

回顾定理1的证明过程，还有这样一句话，即对角矩阵对角线上元素的位置不论怎样掉换，其结果仍与原矩阵相似. 在那里，仅就对角线上前两个元素互换位置来考虑的，一般地说，若 $(a_1, a_2, \cdots, a_n)$ 是 $(b_1, b_2, \cdots, b_n)$ 的一个排列，易证：若

$\boldsymbol{A}=\mathrm{diag}\{a_1, a_2, \cdots, a_n\}$，则

$$\boldsymbol{E}_{ij}^{-1}\boldsymbol{A}\boldsymbol{E}_{ij}=\boldsymbol{E}_{ij}\cdot\mathrm{diag}\{a_1, \cdots, a_i, \cdots, a_j, \cdots, a_n\}\cdot\boldsymbol{E}_{ij}$$
$$=\mathrm{diag}\{a_1, \cdots, a_j, \cdots, a_i, \cdots, a_n\},$$

故有有限多个初等矩阵 $\boldsymbol{P}_1, \cdots, \boldsymbol{P}_k$（每个 $\boldsymbol{P}_i$ 都是 $\boldsymbol{E}_{ij}$ 类型的初等矩阵），使

$$\boldsymbol{P}_k^{-1}\cdots\boldsymbol{P}_2^{-1}\boldsymbol{P}_1^{-1}\boldsymbol{A}\boldsymbol{P}_1\boldsymbol{P}_2\cdots\boldsymbol{P}_k$$
$$=\boldsymbol{B}=\mathrm{diag}\{b_1, b_2, \cdots, b_n\},$$

即

$$\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}=\boldsymbol{B}, \quad \boldsymbol{P}=\boldsymbol{P}_1\boldsymbol{P}_2\cdots\boldsymbol{P}_k.$$

反之，若

$$\boldsymbol{A}=\mathrm{diag}\{a_1, a_2, \cdots, a_n\}$$

与

$$\boldsymbol{B}=\mathrm{diag}\{b_1, b_2, \cdots, b_n\}$$

相似，则 $\boldsymbol{A}$ 与 $\boldsymbol{E}$ 有相同的特征值，因此 $(a_1, a_2, \cdots, a_n)$ 是 $(b_1, b_2, \cdots, b_n)$ 的一个排列.

于是，即证得了：

定理 4 两个对角矩阵 $\mathrm{diag}\{a_1, a_2, \cdots, a_n\}$ 和 $\mathrm{diag}\{b_1, b_2, \cdots, b_n\}$ 相似的充要条件是 $(a_1, a_2, \cdots, a_n)$ 为 $(b_1, b_2, \cdots, b_n)$ 的排列.

能与对角矩阵相似的矩阵，还具有简便的方法去求它的幂. 我们知道，如果根据矩阵乘法的规则，直接去求 $\boldsymbol{A}^k$（当 k 相当大时）是件很麻烦的工作. 但是，如果

$$\boldsymbol{A}\sim\mathrm{diag}\{\lambda_1, \cdots, \lambda_n\},$$

则有满秩矩阵 $\boldsymbol{P}$ 使

$$\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}=\mathrm{diag}\{\lambda_1, \cdots, \lambda_n\},$$

因之

$$\boldsymbol{P}^{-1}\boldsymbol{A}^k\boldsymbol{P}=(\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P})^k$$
$$=\mathrm{diag}\{\lambda_1^k, \lambda_2^k, \cdots, \lambda_n^k\},$$

于是有 $\boldsymbol{A}^k=\boldsymbol{P}\cdot\mathrm{diag}\{\lambda_1^k, \lambda_2^k, \cdots, \lambda_n^k\}\cdot\boldsymbol{P}^{-1}$，以此去求 $\boldsymbol{A}^k$ 一般

说来是较简单的，这时只要去求特征值 $\lambda_1, \lambda_2, \cdots, \lambda_n$ 以及对应于它们的一组 n 个线性无关的特征向量，从而即可求出 $\boldsymbol{P}$. 今举一例来说明.

［例］ 设 $\boldsymbol{A}=\begin{pmatrix}2 & 1\\ 2 & 3\end{pmatrix}$，求 $\boldsymbol{A}^{100}$.

解 因

$$|\lambda \boldsymbol{E}-\boldsymbol{A}|=\begin{vmatrix}\lambda-2 & -1\\ -2 & \lambda-3\end{vmatrix}=(\lambda-1)(\lambda-4),$$

故 $\boldsymbol{A}$ 之两个特征值是 1 与 4，而无重特征值，于是 $\boldsymbol{A}$ 能与对角矩阵相似. 今取分别对应于 1, 4 的一个特征向量，如 $\begin{pmatrix}1\\ -1\end{pmatrix}$ 及 $\begin{pmatrix}1\\ 2\end{pmatrix}$，则 $P=\begin{pmatrix}1 & 1\\ -1 & 2\end{pmatrix}$ 是满秩的，且有

$$\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}=\begin{pmatrix}1 & 0\\ 0 & 4\end{pmatrix},$$

故

$$\begin{aligned}\boldsymbol{A}^{100}&=\boldsymbol{P}\begin{pmatrix}1 & 0\\ 0 & 4^{100}\end{pmatrix}\boldsymbol{P}^{-1}\\ &=\begin{pmatrix}1 & 1\\ -1 & 2\end{pmatrix}\begin{pmatrix}1 & 0\\ 0 & 4^{100}\end{pmatrix}\begin{pmatrix}\frac{2}{3} & -\frac{1}{3}\\ \frac{1}{3} & \frac{1}{3}\end{pmatrix}\\ &=\begin{pmatrix}\frac{2+4^{100}}{3} & \frac{-1+4^{100}}{3}\\ \frac{-2+2\times 4^{100}}{3} & \frac{1+2\times 4^{100}}{3}\end{pmatrix}.\end{aligned}$$

特别地，当 $\boldsymbol{A}$ 之特征值 $\lambda_1, \lambda_2, \cdots, \lambda_n$ 均为 1 之 k 次幂根时，我们还不必去计算 $\boldsymbol{P}$，只要承认 $\boldsymbol{P}$ 之存在，就可以很容

易地知道 $A^k=E$. 这是因为

$$\begin{aligned} P^{-1}A^kP &= (P^{-1}AP)^k \\ &= \mathrm{diag}\{\lambda_1^k, \lambda_2^k, \cdots, \lambda_n^k\} \\ &= E. \end{aligned}$$

练习二十三

1. 若有正整数 k, 使 $A^k=E$, 则 A 相似于对角矩阵.

2. 若 $A\neq 0$, 并有一自然数 $k(\geqslant 2)$ 使 $A^k=0$, 就叫 A 是**幂零矩阵**. 证明: 幂零矩阵不能与对角矩阵相似.

3. 设 A 适合 $A^2+A=2E$, 问 A 能否与对角矩阵相似?

4. 设二级实矩阵 A 的行列式是负数, 试证: 存在二级满秩实矩阵 T, 使 $T^{-1}AT$ 为对角矩阵.

5. 凡 $A^2=A$ 的矩阵 A 叫**幂等矩阵**. 证明: 若幂等矩阵 A 之秩 $=r$, 则 A 与形如 $\begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix}$ 的对角矩阵相似(E_r 为 r 级单位矩阵), 并且 $A+E$ 之行列式为 2^r.

6. 设 $A=\begin{pmatrix} -2 & -1 & -1 & 0 \\ 3 & 1 & 2 & 0 \\ 0 & 1 & -1 & -1 \\ 3 & 0 & 4 & 1 \end{pmatrix}$, 求 A^{60}.

7. 设 n 级矩阵 A 之 n 个特征值两两互异, 证明: 凡具性质 $AB=BA$ 之矩阵 B 必与对角矩阵相似, 且这样的 B 必为且仅为 $B=f(A)$ 形, 其中 $f(x)$ 为 x 之多项式.

8. 任何 n 级矩阵 A 恒可表写为一个满秩矩阵与一个幂等矩阵的积.

9. 设 n 级矩阵 A 的 n 个特征值互异, B 与 A 有相同的特征值, 试证有满秩矩阵 Q 及另一矩阵 R 使得 $A=QR$, $B=RQ$.

第 6 节　矩阵的有理标准形和约旦标准形及其应用

上节研究了矩阵与对角矩阵相似的问题, 并讨论了怎样

的矩阵能与对角矩阵相似．但对于一般不能与对角矩阵相似的矩阵，究竟能与怎样简单形状的矩阵相似呢？ 本节讨论这个问题．

我们知道：$\boldsymbol{A}$ 与 $\boldsymbol{B}$ 相似的充要条件是 $\lambda\boldsymbol{E}-\boldsymbol{A}\simeq\lambda\boldsymbol{E}-\boldsymbol{E}$，故所求条件当然联系上了 $\lambda\boldsymbol{E}-\boldsymbol{A}$ 与 $\lambda\boldsymbol{E}-\boldsymbol{B}$ 有相同的不变因子及初等因子．于是，想解决一个已知矩阵 $\boldsymbol{A}$ 和某个简单形状之矩阵 $\boldsymbol{B}$ 相似的问题，从 $\lambda\boldsymbol{E}-\boldsymbol{A}$ 的不变因子和初等因子去着手是很自然的思想方法．

已知特征矩阵 $\lambda\boldsymbol{E}-\boldsymbol{A}$ 是 λ-矩阵，它的不变因子一般是 λ 的多项式．今反问：任给一多项式 $f(\lambda)$，能否找着一个矩阵（数量矩阵）使 $f(\lambda)$ 是它的特征矩阵的不变因子或初等因子呢？第 2 节末的两个例（参见第 199～200 页）对这个问题作了一个回答．例 2 说明矩阵

$$\boldsymbol{A}=\begin{pmatrix} 0 & 0 & 0 & \cdots & 0 & -a_n \\ 1 & 0 & 0 & \cdots & 0 & -a_{n-1} \\ 0 & 1 & 0 & \cdots & 0 & -a_{n-2} \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & -a_2 \\ 0 & 0 & 0 & \cdots & 1 & -a_1 \end{pmatrix}$$

的特征矩阵

$$\lambda\boldsymbol{E}-\boldsymbol{A}=\begin{pmatrix} \lambda & 0 & 0 & \cdots & 0 & a_n \\ -1 & \lambda & 0 & \cdots & 0 & a_{n-1} \\ 0 & -1 & \lambda & \cdots & 0 & a_{n-2} \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda & a_2 \\ 0 & 0 & 0 & \cdots & -1 & \lambda+a_1 \end{pmatrix}$$

只有唯一个非常数的不变因子

$$f(\lambda)=\lambda^n+a_1\lambda^{n-1}+\cdots+a_{n-1}\lambda+a_n,$$

因之

$$\lambda E-A\simeq\begin{pmatrix}1&&&\\&\ddots&&\\&&1&\\&&&f(\lambda)\end{pmatrix},$$

这说明了给了一个 n 次多项式 $f(\lambda)$，确有一个 n 级矩阵，它的特征矩阵只有唯一个非常数的不变因子 $f(\lambda)$. 于是我们得知给了一个矩阵，对于它的特征矩阵的每一个非常数的不变因子，都可以求得象刚才说的那样的矩阵，这就提供了从不变因子着手去求出与已知矩阵相似而形状又较简单的矩阵，具体求法如下：

设 $A=(a_{ij})$ 是 n 级矩阵，其特征矩阵 $\lambda E-A$ 中非常数的不变因子有 s 个：$f_1(\lambda)$, …, $f_s(\lambda)$，因之有 $f_i(\lambda)\mid f_{i+1}(\lambda)$ $(i=1, 2, \cdots, s-1)$. 于是，

$$\lambda E-A\simeq\begin{pmatrix}1&&&&&\\&\ddots&&&&\\&&1&&&\\&&&f_1(\lambda)&&\\&&&&\ddots&\\&&&&&f_s(\lambda)\end{pmatrix},$$

命

$$f_i(\lambda)=\lambda^{n_i}+a_{i1}\lambda^{n_i-1}+a_{i2}\lambda^{n_i-2}+\cdots+a_{i,n_i-1}\lambda+a_{in_i},$$

并作 n 级矩阵

$$F=\begin{pmatrix}N_1&&&\\&N_2&&\\&&\ddots&\\&&&N_s\end{pmatrix},$$

其中

$$
N_i=\begin{pmatrix} 0 & 0 & 0 & \cdots & 0 & -a_{in_i} \\ 1 & 0 & 0 & \cdots & 0 & -a_{i,n_i-1} \\ 0 & 1 & 0 & \cdots & 0 & -a_{i,n_i-2} \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & -a_{i2} \\ 0 & 0 & 0 & \cdots & 1 & -a_{i1} \end{pmatrix}
$$

为 n_i 级矩阵，这里 $n_1+n_2+\cdots+n_s=n$.

用这样的方法作出的 n 级矩阵 $\boldsymbol{F}$ 是与 $\lambda\boldsymbol{E}-\boldsymbol{A}$ 中的不变因子紧密地联系着的. 显然，因 $\boldsymbol{F}$ 是分块矩阵，故一般地讲，形状是较简单的，若能证明 $\boldsymbol{A}\sim\boldsymbol{F}$，则 $\boldsymbol{F}$ 即为我们所要求形状较简单而又与 $\boldsymbol{A}$ 相似的一个矩阵，且是从 $\lambda\boldsymbol{E}-\boldsymbol{A}$ 之不变因子着手而得到的.

定理 1 $\boldsymbol{A}\sim\boldsymbol{F}$.

通常，叫 $\boldsymbol{F}$ 为 $\boldsymbol{A}$ 之*有理标准形*，之所以这样叫，其原因不外乎是 $\boldsymbol{F}$ 的元素是从 $\boldsymbol{A}$ 的元素经过有理运算(即加减乘除四则运算)求得的.

证明 欲证 $\boldsymbol{A}\sim\boldsymbol{F}$，就只需证明 $\lambda\boldsymbol{E}-\boldsymbol{A}\simeq\lambda\boldsymbol{E}-\boldsymbol{F}$ 就行了(第 3 节定理 2). 事实上，

$$
\lambda\boldsymbol{E}-\boldsymbol{F}=\begin{pmatrix} \lambda\boldsymbol{E}-\boldsymbol{N}_1 & & & \\ & \lambda\boldsymbol{E}-\boldsymbol{N}_2 & & \\ & & \ddots & \\ & & & \lambda\boldsymbol{E}-\boldsymbol{N}_s \end{pmatrix},
$$

这里的 $\lambda\boldsymbol{E}-\boldsymbol{F}$ 中之 $\boldsymbol{E}$ 是 n 级单位矩阵，而 $\lambda\boldsymbol{E}-\boldsymbol{N}_i$ 中之 $\boldsymbol{E}$ 是 n_i 级单位矩阵，为简便而用同一个符号. 但

$$
\lambda\boldsymbol{E}-\boldsymbol{N}_i\simeq\begin{pmatrix} 1 & & & \\ & \ddots & & \\ & & 1 & \\ & & & f_i(\lambda) \end{pmatrix},
$$

所以

$$\lambda\boldsymbol{E}-\boldsymbol{F}\simeq\begin{pmatrix}1&&&&&&&\\&\ddots&&&&&&\\&&1&&&&&\\&&&f_1(\lambda)&&&&\\&&&&\ddots&&&\\&&&&&1&&\\&&&&&&\ddots&\\&&&&&&1&\\&&&&&&&f_s(\lambda)\end{pmatrix}$$

$$\simeq\begin{pmatrix}1&&&&&\\&\ddots&&&&\\&&1&&&\\&&&f_1(\lambda)&&\\&&&&\ddots&\\&&&&&f_s(\lambda)\end{pmatrix}.$$

这就是说 $\lambda\boldsymbol{E}-\boldsymbol{F}$ 和 $\lambda\boldsymbol{E}-\boldsymbol{A}$ 有相同的不变因子，因此它们等价($\lambda\boldsymbol{E}-\boldsymbol{F}\simeq\lambda\boldsymbol{E}-\boldsymbol{A}$)，即定理 1 得证.

当 $\boldsymbol{A}$ 之特征矩阵 $\lambda\boldsymbol{E}-\boldsymbol{A}$ 中 n 个不变因子都是非常数的时候，显然 $I_1(\lambda)$, …, $I_n(\lambda)$ 都是一次式，因此，

$$I_1(\lambda)=I_2(\lambda)=\cdots=I_n(\lambda)=\lambda-a,$$

这时 $\boldsymbol{F}$ 中的各个 $\boldsymbol{N}_i=a$，于是

$$\boldsymbol{A}\sim\begin{pmatrix}a&&&\\&a&&\\&&\ddots&\\&&&a\end{pmatrix}=\operatorname{diag}\{a,\ a,\ \cdots,\ a\}=a\boldsymbol{E},$$

即 $\boldsymbol{A}$ 与纯量矩阵相似，因此 $\boldsymbol{A}=\boldsymbol{P}^{-1}(a\boldsymbol{E})\boldsymbol{P}=a\boldsymbol{E}$. 这就是说 $\boldsymbol{A}$ 也是纯量矩阵. 反之，纯量矩阵的不变因子显然都为非常数的，实际上都是 λ 的一次式. 故有：

推论 矩阵 $\boldsymbol{A}$ 是纯量矩阵的充要条件是它的特征矩阵的不变因子都非常数.

下面来谈另一个相似矩阵的形状，它是与它的特征矩阵的初等因子有连系的.

第 2 节的例 1 (参见第 199 页) 又证实了矩阵

$$\boldsymbol{A}=\begin{pmatrix} a & 0 & 0 & \cdots & 0 & 0 \\ 1 & a & 0 & \cdots & 0 & 0 \\ 0 & 1 & a & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & a & 0 \\ 0 & 0 & 0 & \cdots & 1 & a \end{pmatrix}$$

的特征矩阵

$$\lambda\boldsymbol{E}-\boldsymbol{A}=\begin{pmatrix} \lambda-a & 0 & 0 & \cdots & 0 & 0 \\ -1 & \lambda-a & 0 & \cdots & 0 & 0 \\ 0 & -1 & \lambda-a & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda-a & 0 \\ 0 & 0 & 0 & \cdots & -1 & \lambda-a \end{pmatrix}$$

也只有唯一个非常数的不变因子 $f(\lambda)=(\lambda-a)^n$，因而它也只有一个初等因子 $(\lambda-a)^n$，这时我们有

$$\lambda\boldsymbol{E}-\boldsymbol{A}\simeq\begin{pmatrix} 1 & & & \\ & \ddots & & \\ & & 1 & \\ & & & (\lambda-a)^n \end{pmatrix}.$$

因此，给了一个 n 次多项式是一次因式之 n 次幂 (如 $(\lambda-a)^n$) 时，就可作一个 n 级矩阵，使它的特征矩阵只有 $(\lambda-a)^n$ 为它的唯一个初等因子. 于是，给了一个矩阵，对于它的特征矩阵之每一个初等因子，都可以求得象刚才说的这样的矩阵，这就提供了从初等因子着手去求形状较简单之矩阵使与已给矩阵相似的问题，具体求法于下：

设 $\boldsymbol{A}=(a_{ij})$ 是 n 级矩阵，它的特征矩阵 $\lambda\boldsymbol{E}-\boldsymbol{A}$ 的初等因子是

$$(\lambda-\lambda_1)^{k_1},\ (\lambda-\lambda_2)^{k_2},\ \cdots,\ (\lambda-\lambda_t)^{k_t},$$

这里，当 $i\neq j$ 时可能有 $\lambda_i=\lambda_j$，显然 $k_1+k_2+\cdots+k_t=n$. 作 k_i 级矩阵

$$\boldsymbol{J}_i=\begin{pmatrix}\lambda_i & 0 & 0 & \cdots & 0 & 0\\ 1 & \lambda_i & 0 & \cdots & 0 & 0\\ 0 & 1 & \lambda_i & \cdots & 0 & 0\\ \vdots & \vdots & \vdots & & \vdots & \vdots\\ 0 & 0 & 0 & \cdots & \lambda_i & 0\\ 0 & 0 & 0 & \cdots & 1 & \lambda_i\end{pmatrix},$$

并作 n 级矩阵

$$\boldsymbol{J}=\begin{pmatrix}\boldsymbol{J}_1 & & & \\ & \boldsymbol{J}_2 & & \\ & & \ddots & \\ & & & \boldsymbol{J}_t\end{pmatrix}.$$

若能证明 $\boldsymbol{A}\sim\boldsymbol{J}$，则因 $\boldsymbol{J}$ 之形状当然较 $\boldsymbol{A}$ 简单些，故 $\boldsymbol{J}$ 即为我们所要求的从 $\lambda\boldsymbol{E}-\boldsymbol{A}$ 之初等因子出发而得到与 $\boldsymbol{A}$ 相似且形状又简单些的一个矩阵.

定理 2 $\boldsymbol{A}\sim\boldsymbol{J}$.

通常叫 $\boldsymbol{J}$ 为 $\boldsymbol{A}$ 之约旦标准形，是约旦(Jordan)首先提出的，每个 $\boldsymbol{J}_i$ 叫做约旦矩阵，并叫做 $\boldsymbol{A}$(或 $\boldsymbol{J}$)的约旦块.

证明 因为

$$\lambda\boldsymbol{E}-\boldsymbol{J}_i\simeq\begin{pmatrix}1 & & & \\ & \ddots & & \\ & & 1 & \\ & & & (\lambda-\lambda_i)^{k_i}\end{pmatrix},$$

所以

$$\lambda E-J=\begin{pmatrix}\lambda E-J_1 & & & \\ & \lambda E-J_2 & & \\ & & \ddots & \\ & & & \lambda E-J_t\end{pmatrix}$$

$$\simeq\begin{pmatrix}1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & & (\lambda-\lambda_1)^{k_1} & & \\ & & & & \ddots & \\ & & & & & (\lambda-\lambda_t)^{k_t}\end{pmatrix}.$$

根据第 2 节的定理 7，得知$(\lambda-\lambda_1)^{k_1}$，…，$(\lambda-\lambda_t)^{k_t}$是$\lambda E-J$的初等因子，因此$\lambda E-A\simeq\lambda E-J$. 定理 2 获证.

当A之特征矩阵$\lambda E-A$的初等因子都是一次式时，即$k_1=k_2=\cdots=k_t=1$，这时当然$t=n$. 于是，每个J_i都是一级的，即$J_i=\lambda_i$. 因此，就有$J=\mathrm{diag}\{\lambda_1, \cdots, \lambda_n\}$. 这就是第 5 节定理 2 中条件之充分性的另一证明.

再来谈谈有理标准形和约旦标准形在解微分方程组时的应用.

(一)利用约旦标准形，解

$$\frac{dx}{dt}=Ax, \tag{21}$$

这里$A=(a_{ij})$是n级矩阵，$x=\begin{pmatrix}x_1\\ \vdots\\ x_n\end{pmatrix}$，$\dfrac{dx}{dt}=\begin{pmatrix}\dfrac{dx_1}{dt}\\ \vdots\\ \dfrac{dx_n}{dt}\end{pmatrix}$. 先把$A$化为约旦标准形$J$，$J=PAP^{-1}$，式中$P=(p_{ij})$. 由(21)即得

$$P\cdot\frac{dx}{dt}=PAx=PAP^{-1}\cdot Px.$$

但

$$\boldsymbol{P}\cdot\frac{d\boldsymbol{x}}{dt}=(p_{ij})\cdot\begin{pmatrix}\frac{dx_1}{dt}\\ \vdots\\ \frac{dx_n}{dt}\end{pmatrix}=\begin{pmatrix}p_{11}\frac{dx_1}{dt}+\cdots+p_{1n}\frac{dx_n}{dt}\\ \vdots\\ p_{n1}\frac{dx_1}{dt}+\cdots+p_{nn}\frac{dx_n}{dt}\end{pmatrix}$$

$$=\begin{pmatrix}\frac{d}{dt}(p_{11}x_1+\cdots+p_{1n}x_n)\\ \vdots\\ \frac{d}{dt}(p_{n1}x_1+\cdots+p_{nn}x_n)\end{pmatrix}=\frac{d}{dt}\boldsymbol{P}\boldsymbol{x}.$$

因此有

$$\frac{d\boldsymbol{y}}{dt}=\boldsymbol{J}\boldsymbol{y},\quad \boldsymbol{y}=\begin{pmatrix}y_1\\ \vdots\\ y_n\end{pmatrix}=\boldsymbol{P}\boldsymbol{x}=\boldsymbol{P}\begin{pmatrix}x_1\\ \vdots\\ x_n\end{pmatrix},$$

$$\frac{d\boldsymbol{y}}{dt}=\begin{pmatrix}\frac{dy_1}{dt}\\ \vdots\\ \frac{dy_n}{dt}\end{pmatrix}$$

$$=\begin{pmatrix}\boxed{\begin{matrix}\lambda_1 & 0 & \cdots & 0 & 0\\ 1 & \lambda_1 & \cdots & 0 & 0\\ \vdots & \vdots & & \vdots & \vdots\\ 0 & 0 & \cdots & 1 & \lambda_1\end{matrix}} & & \\ & \ddots & \\ & & \boxed{\begin{matrix}\lambda_t & 0 & \cdots & 0 & 0\\ 1 & \lambda_t & \cdots & 0 & 0\\ \vdots & \vdots & & \vdots & \vdots\\ 0 & 0 & \cdots & 1 & \lambda_t\end{matrix}}\end{pmatrix}\begin{pmatrix}y_1\\ y_2\\ \vdots\\ y_n\end{pmatrix}.$$

于是，

$$\frac{dy_1}{dt}=\lambda_1 y_1,\ \frac{dy_2}{dt}=y_1+\lambda_1 y_2,\ \cdots,\ \frac{dy_{k_1}}{dt}=y_{k_1-1}+\lambda_1 y_{k_1};$$

$$\cdots\cdots\cdots\cdots\cdots\cdots$$

$$\frac{dy_{k_1+\cdots+k_{t-1}+1}}{dt}=\lambda_t y_{k_1+\cdots+k_{t-1}+1},\ \cdots,\ \frac{dy_n}{dt}=y_{n-1}+\lambda_t y_n.$$

从 $\frac{dy_1}{dt}=\lambda_1 y_1$ 得 $y_1=c_1e^{\lambda_1 t}$（c_1 是参数）．再令 $y_2=e^{\lambda_1 t}v$（v 是 t 的函数），从 $\frac{dy_2}{dt}=y_1+\lambda_1 y_2$ 可得 $y_2=e^{\lambda_1 t}(c_1t+c_2)$，$c_2$ 是参数．同样，令 $y_3=e^{\lambda_1 t}u$（其中 u 是 t 的函数）又得

$$y_3=e^{\lambda_1 t}\left(\frac{c_1}{2!}t^2+c_2t+c_3\right).$$

这样继续往下做，就得到

$$y_{k_1}=e^{\lambda_1 t}\left[\frac{c_1}{(k_1-1)!}t^{k_1-1}+\frac{c_2}{(k_1-2)!}t^{k_1-2}+\cdots\right.$$
$$\left.+\frac{c_{k_1-2}}{2!}t^2+c_{k_1-1}t+c_{k_1}\right],$$

其中 $c_1,\ c_2,\ \cdots,\ c_{k_1}$ 都是独立参数．

用类似的方法，又可将 $y_{k_1+1},y_{k_1+2},\cdots,y_{k_1+k_2}$ 顺次地求出，继续往下做，一直可以把 $y_{k_1+\cdots+k_{t-1}+1},\ \cdots,\ y_n$ 顺次地求出来．

求得 $y_1,\ \cdots,\ y_n$ 后，由 $\boldsymbol{x}=\boldsymbol{P}^{-1}\boldsymbol{y}$ 即得 $x_1,\ \cdots,\ x_n$，即为所求线性微分方程组(21)之解．

（二）利用有理标准形来解方程组(21)．将 $\boldsymbol{A}$ 化为有理标准形 $\boldsymbol{F}$，令 $\boldsymbol{F}=\boldsymbol{PAP}^{-1}$．同上面一样，令 $\boldsymbol{y}=\boldsymbol{Px}$，即得

$$\frac{d\boldsymbol{y}}{dt}=\begin{pmatrix}\frac{dy_1}{dt}\\ \vdots\\ \frac{dy_n}{dt}\end{pmatrix}=\begin{pmatrix}\boldsymbol{N}_1&&&\\&\boldsymbol{N}_2&&\\&&\ddots&\\&&&\boldsymbol{N}_s\end{pmatrix}\begin{pmatrix}y_1\\ y_2\\ \vdots\\ y_n\end{pmatrix},$$

对于 N_1，有

$$\begin{pmatrix}\frac{dy_1}{dt}\\ \vdots \\ \frac{dy_{n_1}}{dt}\end{pmatrix}=\begin{pmatrix}0&0&0&\cdots&0&-a_{1n_1}\\1&0&0&\cdots&0&-a_{1,n_1-1}\\0&1&0&\cdots&0&-a_{1,n_1-2}\\ \vdots&\vdots&\vdots&&\vdots&\vdots\\0&0&0&\cdots&0&-a_{12}\\0&0&0&\cdots&1&-a_{11}\end{pmatrix}\begin{pmatrix}y_1\\y_2\\ \vdots\\y_{n_1}\end{pmatrix}.$$

于是，

$$\frac{dy_1}{dt}=-a_{1n_1}y_{n_1},\ \frac{dy_2}{dt}=y_1-a_{1,n_1-1}y_{n_1},\ \cdots,\ \frac{dy_{n_1}}{dt}=y_{n_1-1}-a_{11}y_{n_1}.$$

因此，y_1 可表为 y_{n_1} 的积分，从而 y_2 也可表为 y_{n_1} 的积分，继续下推，可知 y_3，…，y_{n_1-1} 都可用 y_{n_1} 的积分式来表示．由最后的表示式再先求出 y_{n_1}，然后顺次就可求得 y_1，y_2，…，y_{n_1-1}.

对于其他的 N_i，可以同样处理．在求得 y_1，y_2，…，y_n 之后，再利用 $\boldsymbol{x}=\boldsymbol{P}^{-1}\boldsymbol{y}$，就可以把 x_1，x_2，…，x_n 全部求出．

这样，利用约旦标准形或有理标准形，就可以把解微分方程组的手续大为简化．比较这两种方法，各有短长．有了约旦标准形，y_1，y_2，…，y_n 就可以直接写出，因此很易求得 x_1，x_2，…，x_n，这是它的优点；但求 $\boldsymbol{A}$ 的特征值 λ_1，…，λ_n 不是一件容易事，在实际计算中有极大的困难；利用有理标准形可以避免求特征值，但在求 y_1，…，y_n 时需要求积分也非易事．至于用哪一方法适宜，应由方程组(21)的实际情况来决定．

练习二十四

1．设 $\boldsymbol{A}$ 为幂零矩阵，试求 $\boldsymbol{A}+\boldsymbol{E}$ 之行列式．

2．若幂零矩阵 $\boldsymbol{A}$ 的秩为 r，试证 $\boldsymbol{A}^r=\boldsymbol{O}$．

3．在 $\mathbb{F}[x]_n$ 中，求微商变换 $\mathscr{D}$ 的特征多项式，并证明在任意一组基底下的 $\mathscr{D}$ 之矩阵 $\boldsymbol{D}$ 都不与对角矩阵相似．

4. 化$A=\begin{pmatrix}0&1&1\\1&0&1\\1&1&0\end{pmatrix}$为有理标准形。

5. 化$A=\begin{pmatrix}2&-1&1\\2&2&-1\\1&2&-1\end{pmatrix}$为约旦标准形。

第7节 矩阵之标准形的几何意义

对于矩阵的有理标准形，约旦标准形，及其特例的对角形，现在来说明这些标准形在几何学上的解释．

先从有理标准形来谈起．为便于说明，应注意矩阵$\boldsymbol{A}$的有理标准形一般有两种表示法，前节的定理1只是其中的一种，另一种是

$$\boldsymbol{A}\sim\boldsymbol{F}',$$

$\boldsymbol{F}'$是前节定理1中$\boldsymbol{F}$的转置矩阵，即

$$\boldsymbol{F}'=\begin{pmatrix}\boldsymbol{N}_1' & & & 0\\ & \boldsymbol{N}_2' & & \\ & & \ddots & \\ 0 & & & \boldsymbol{N}_s'\end{pmatrix},$$

而

$$\boldsymbol{N}_i'=\begin{pmatrix}0&1&0&\cdots&0&0\\0&0&1&\cdots&0&0\\0&0&0&\cdots&0&0\\\vdots&\vdots&\vdots&&1&\vdots\\0&0&0&\cdots&0&1\\-a_{in_i}&-a_{i,n_i-1}&-a_{i,n_i-2}&\cdots&-a_{i2}&-a_{i1}\end{pmatrix}.$$

证明的方法完全一样．采用这一种表示法的目的，是为了便

于说明几何意义．例如，设

$$N_1'=\begin{pmatrix} 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 \\ 0 & 0 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & 1 \\ a_1 & a_2 & a_3 & \cdots & a_{k-1} & a_k \end{pmatrix}$$

为 k 级矩阵，则

$$A\sim F'=\begin{pmatrix} N_1' & & & \\ & N_2' & & \\ & & \ddots & \\ & & & N_s' \end{pmatrix}$$

是说明：在 n 维空间 $\mathfrak{R}_n$ 中，A 与 F' 对应于同一个线性变换 $\mathscr{T}$，亦即有这样一组基底 $e_1, e_2, \cdots, e_n$，使其前 k 个向量 $e_1, e_2, \cdots, e_k$ 具有

$$\begin{cases} e_1\mathscr{T}=e_2, \\ e_2\mathscr{T}=e_3, \\ \cdots\cdots\cdots\cdots \\ e_{k-1}\mathscr{T}=e_k, \\ e_k\mathscr{T}=a_1e_1+\cdots+a_ke_k \end{cases} \tag{22}$$

之关系．同样，后 $n-k$ 个向量 $e_{k+1}, \cdots, e_n$ 可分为 $s-1$ 组，使每组均有类似于(22)之关系．

(22)式的几何意义是：由 $e_1, \cdots, e_k$ 为基底所生成的 $\mathfrak{R}_n$ 之 k 维子空间 $\mathfrak{S}_1=\mathfrak{S}[e_1, e_2, \cdots, e_k]$ 对于线性变换 $\mathscr{T}$ 是不变的，即 $\mathscr{T}$ 使 $\mathfrak{S}_1$ 的每个向量变化的结果仍然在 $\mathfrak{S}_1$ 内($\mathfrak{S}_1\mathscr{T}\sqsubseteq\mathfrak{S}_1$)，象这样的子空间 $\mathfrak{S}_1$ 叫做线性变换 $\mathscr{T}$ 的不变子空间．于是，全空间 $\mathfrak{R}_n$ 可分解为 s 个不变子空间的和．又，这 s 个子空间中任两个的交为零(即零空间)，且 $\mathfrak{R}_n$ 的每个向量表为

从这 s 个子空间内各取一向量之和的表示方法也只有一种，我们把象这样的和叫做直和，于是 $\mathfrak{R}_n$ 可以分解为 s 个不变子空间的直和(不变的意义指对线性变换 $\mathscr{T}$ 而言)．

又，(22)式同时也说明了 $e_1, e_2, e_3, \cdots, e_{k-1}, e_k$ 分别等于 $e_1\mathscr{T}^0, e_1\mathscr{T}, e_1\mathscr{T}^2, \cdots, e_1\mathscr{T}^{k-2}, e_1\mathscr{T}^{k-1}$(其中 $\mathscr{T}^0$ 表示 $\mathscr{T}$ 的零次幂，即为恒等变换)，即不变子空间 $\mathfrak{S}_1$ 的一组基底 $e_1, e_2, \cdots, e_k$ 可以由 $\mathscr{T}^0, \mathscr{T}, \mathscr{T}^2, \cdots, \mathscr{T}^{k-1}$ 分别作用于同一个向量 e_1 而得到，即 $\mathfrak{S}_1$ 之一组基底可由 $\mathscr{T}$ 之诸幂作用于同一向量而产生．具有这样一组基底的不变子空间叫做循环不变子空间．

再注意 N_1' 之特征矩阵 $\lambda E-N_1'$ 只有一个非常数的不变因子，就知道它的最小多项式即为其特征多项式，因而 N_i' 之最小多项式为 N_{i+1}' 之最小多项式的因式．但 $\mathscr{T}$ 在空间 $\mathfrak{S}_1$ 内对应的矩阵均与 N_1' 相似，而相似矩阵有相同的最小多项式，故 N_1' 之最小多项式与 $\mathfrak{S}_1$ 的基底之选择无关，而仅与线性变换 $\mathscr{T}$ 自身有关，因之把 N_1' 的最小多项式就叫做线性变换 $\mathscr{T}$ 的最小多项式，是不会引起误解的．

又，由于矩阵 A 之特征矩阵的不变因子是由 A 唯一确定的，即有理标准形 F' 由 A 唯一确定，随而 $N_1', \cdots, N_s'$ 亦然，因之循环不变子空间 $\mathfrak{S}_1, \mathfrak{S}_2, \cdots, \mathfrak{S}_s$ 除同构外应由 $\mathscr{T}$(即 A)唯一确定．

总括上述，可知任一矩阵必有有理标准形的几何意义即是：

定理 1 对于 n 维空间 $\mathfrak{R}_n$ 的任一个线性变换 $\mathscr{T}$，我们总可以把 $\mathfrak{R}_n$ 分解为 s 个($s\geqslant 1$，s 由 $\mathscr{T}$ 而定)这样的循环不变子空间 $\mathfrak{S}_1, \mathfrak{S}_2, \cdots, \mathfrak{S}_s$ 的直和，表为

$$\mathfrak{R}_n=\mathfrak{S}_1\oplus\mathfrak{S}_2\oplus\cdots\oplus\mathfrak{S}_s,$$

使得 $\mathscr{T}$ 在 $\mathfrak{S}_i$ 内的最小多项式为它在 $\mathfrak{S}_{i+1}$ 内的最小多项式的因式．并且这样的分解除同构的外是唯一的†．

［例 1］ 试讨论三维空间中的向量投影到坐标面 XOY 上的投影变换．

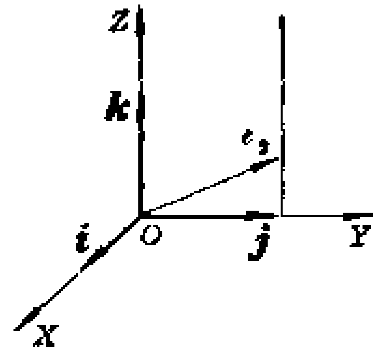


图 5-1

解 令 $\mathscr{T}$ 表示所说的投影变换．因 $\boldsymbol{i}\mathscr{T}=\boldsymbol{i}$, $\boldsymbol{j}\mathscr{T}=\boldsymbol{j}$, $\boldsymbol{k}\mathscr{T}=\boldsymbol{0}$, 故 $\mathscr{T}$ 对应的矩阵(对基底 $\boldsymbol{i}$, $\boldsymbol{j}$, $\boldsymbol{k}$ 面言)为

$$\boldsymbol{A}=\begin{pmatrix}1&0&0\\0&1&0\\0&0&0\end{pmatrix},$$

于是易知特征矩阵 $\lambda\boldsymbol{E}-\boldsymbol{A}$ 的三个不变因子为 $I_1(\lambda)=1$, $I_2(\lambda)=\lambda-1$, $I_3(\lambda)=\lambda^2-\lambda$, 故 $\boldsymbol{A}$ 之有理标准形为

$$\boldsymbol{F}=\begin{pmatrix}\boxed{1}&\\&\boxed{\begin{matrix}0&1\\0&1\end{matrix}}\end{pmatrix},$$

因面 $\mathfrak{R}_3$ 必有这样一组基底 $\boldsymbol{e}_1$, $\boldsymbol{e}_2$, $\boldsymbol{e}_3$, 使得 $\mathfrak{S}_1=\mathfrak{S}[\boldsymbol{e}_1]$ 与 $\mathfrak{S}_2=\mathfrak{S}[\boldsymbol{e}_2,\ \boldsymbol{e}_3]$ 为两个不变子空间，维数分别等于 1 与 2, 且有 $\mathfrak{R}_3=\mathfrak{S}_1\oplus\mathfrak{S}_2$. 实际上，有 $\boldsymbol{e}_1\mathscr{T}=\boldsymbol{e}_1$, $\boldsymbol{e}_2\mathscr{T}=\boldsymbol{e}_3$, $\boldsymbol{e}_3\mathscr{T}=\boldsymbol{e}_3$. 例如取 $\boldsymbol{e}_1=\boldsymbol{i}$, $\boldsymbol{e}_3=\boldsymbol{j}$, $\boldsymbol{e}_2$ 为在坐标面 YOZ 上而终点在过 $\boldsymbol{j}$ 之終点且平行于 Z 轴的直线上之任意向量(如图 5-1), 这时 $\mathfrak{S}_1$ 为 X-轴, $\mathfrak{S}_2$ 为平面 YOZ. 同理, $\mathfrak{S}_1$ 与 $\mathfrak{S}_2$ 亦可分别为 Y-轴与平面 XOZ. 同样, X-轴与过 Z-轴且垂直于坐标面 XOY

---

† 两个不变子空间 $\mathfrak{S}$ 与 $\mathfrak{S}'$ 同构，指的是：

1) $\mathfrak{S}$ 与 $\mathfrak{S}'$ 在第三章第 5 节定义 3 的意义下同构，因之它们的维数相等．

2) 若在 1) 之同构关系中 $\mathfrak{S}$ 之向量 $\boldsymbol{x}$ 对应于 $\mathfrak{S}'$ 之向量 $\boldsymbol{x}'$ (即 $\boldsymbol{x}\leftrightarrow\boldsymbol{x}'$), 那末 $\boldsymbol{x}\mathscr{T}\leftrightarrow\boldsymbol{x}'\mathscr{T}$.

的任一平面(但不含 X-轴)也可以分别作为 $\mathfrak{S}_1$ 与 $\mathfrak{S}_2$, 等等.

[例 2] 讨论平面中的旋转变换(旋转角为 θ).

解 如图 5-2, 若令 $\mathscr{T}$ 表示平面中的旋转变换(旋转角为 θ), 则由

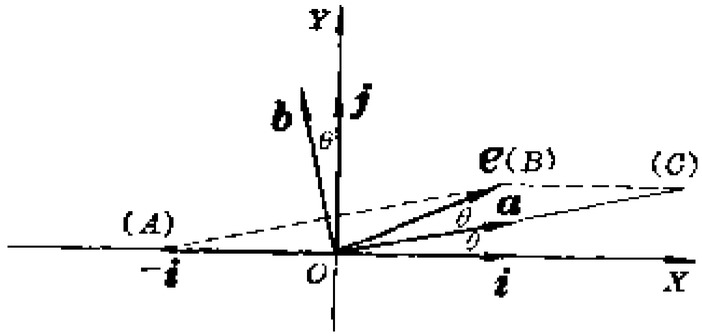


图 5-2

$$\boldsymbol{i}\mathscr{T}=\boldsymbol{a}=\cos\theta\cdot\boldsymbol{i}+\sin\theta\cdot\boldsymbol{j},$$

$$\boldsymbol{j}\mathscr{T}=\boldsymbol{b}=-\sin\theta\cdot\boldsymbol{i}+\cos\theta\cdot\boldsymbol{j},$$

可知: $\mathscr{T}$ 对应之矩阵(对基底 $\boldsymbol{i}, \boldsymbol{j}$)为

$$\boldsymbol{A}=\begin{pmatrix}\cos\theta & \sin\theta\\ -\sin\theta & \cos\theta\end{pmatrix},$$

其特征矩阵 $\lambda\boldsymbol{E}-\boldsymbol{A}$ 的两个不变因子为 $I_1(\lambda)=1$, $I_2(\lambda)=\lambda^2-2\cos\theta\cdot\lambda+1$, 故 $\boldsymbol{A}$ 之有理标准形为 $\boldsymbol{F}=\begin{pmatrix}0 & 1\\ -1 & 2\cdot\cos\theta\end{pmatrix}$. 这时, 不变子空间就是平面自身(因 $\lambda\boldsymbol{E}-\boldsymbol{A}$ 只有一个非常数的不变因子); 且存在一组基底 $\boldsymbol{e}_1$、$\boldsymbol{e}_2$, 使 $\mathscr{T}$ 对应之矩阵即为 $\boldsymbol{F}$, 但 $\boldsymbol{e}_1\mathscr{T}=\boldsymbol{e}_2$, $\boldsymbol{e}_2\mathscr{T}=-\boldsymbol{e}_1+2\cos\theta\cdot\boldsymbol{e}_2$. 例如取 $\boldsymbol{e}_1=\boldsymbol{i}$ 时, 则如图 5-2, $\boldsymbol{e}_2=\boldsymbol{a}$, 这时 $\boldsymbol{e}_2\mathscr{T}=\boldsymbol{a}\mathscr{T}=\boldsymbol{e}$, 且确有

$$\boldsymbol{e}=-\boldsymbol{i}+2\cos\theta\cdot\boldsymbol{a}.$$

事实上, $\triangle OAB$ 是等腰的($OA=OB$), 因向量 $-\boldsymbol{i}$ 与 $\boldsymbol{e}$ 之长均为1, 而各为 OA 与 OB; 又向量 $\boldsymbol{a}$ 平分 $\angle XOB$, 故 $AB\parallel\boldsymbol{a}$. 于是过 B 点作平行于 X-轴之直线而交向量 $\boldsymbol{a}$ 于 C 点时, 则得平行四边形 $ABCO$, 因而

$$\boldsymbol{e}=-\boldsymbol{i}+\overrightarrow{OC}=-\boldsymbol{i}+x\boldsymbol{a},$$

但 x 为向量 $\overrightarrow{OC}$ 之长(因为 $\boldsymbol{a}$ 是单位向量). 然从 $\triangle OAB$ 看, 由余弦定理就有

$$x^2=\overline{AB}^2=\overline{OA}^2+\overline{OB}^2-2\cdot\overline{OA}\cdot\overline{OB}\cdot\cos\angle AOB$$
$$=1+1-2\cos(\pi-2\theta)=2+2\cos 2\theta$$
$$=2(1+\cos 2\theta)=4\cos^2\theta,$$
即 $x=2\cos\theta$(这时为方便计是取 θ 为锐角来讨论的). 故
$$\boldsymbol{e}=-\boldsymbol{i}+2\cos\theta\cdot\boldsymbol{a}.$$

再来谈约旦标准形. 从前节定理 2 之约旦标准形
$$\boldsymbol{J}=\begin{pmatrix}\boldsymbol{J}_1 & & & \\ & \boldsymbol{J}_2 & & \\ & & \ddots & \\ & & & \boldsymbol{J}_t\end{pmatrix},$$
其中
$$\boldsymbol{J}_1=\begin{pmatrix}\lambda_1 & & & & \\ 1 & \lambda_1 & & & \\ & \ddots & \ddots & & \\ & & 1 & \lambda_1 & \\ & & & \ddots & \ddots & \\ & & & & 1 & \lambda_1\end{pmatrix}$$
(k_1 级矩阵)来看即知在 n 维空间 $\mathfrak{R}_n$ 中有这样一组基底 $\boldsymbol{e}_1, \cdots, \boldsymbol{e}_n$, 使前 k_1 个向量 $\boldsymbol{e}_1, \cdots, \boldsymbol{e}_{k_1}$ 生成一个 k_1 维的不变子空间 $\mathfrak{S}_1=\mathfrak{S}[\boldsymbol{e}_1, \cdots, \boldsymbol{e}_{k_1}]$, 而实际上有
$$\begin{cases}\boldsymbol{e}_1\mathscr{T}=\lambda_1\boldsymbol{e}_1,\\ \boldsymbol{e}_2\mathscr{T}=\boldsymbol{e}_1+\lambda_1\boldsymbol{e}_2,\\ \boldsymbol{e}_3\mathscr{T}=\boldsymbol{e}_2+\lambda_1\boldsymbol{e}_3,\\ \cdots\cdots\cdots\cdots\cdots\cdots\\ \boldsymbol{e}_{k_1}\mathscr{T}=\boldsymbol{e}_{k_1-1}+\lambda_1\boldsymbol{e}_{k_1},\end{cases}\tag{23}$$
其中 $\mathscr{T}$ 为矩阵 $\boldsymbol{A}$ 所对应的线性变换.

同样, 对于后 $n-k_1$ 个向量 $\boldsymbol{e}_{k_1+1}, \cdots, \boldsymbol{e}_n$, 可分为 $t-1$ 组, 使每组有类似于(23)之关系, 于是 $\mathfrak{R}_n$ 可分解为 t 个不变子空间的直和; 且又因 $\lambda\boldsymbol{E}-\boldsymbol{A}$ 的初等因子由 $\boldsymbol{A}$ 唯一确定, 故这

样的分解除同构者外是唯一的．于是又有：

定理2 对于 n 维空间 $\Re_n$ 之任一个线性变换 $\mathscr{T}$，总可以把 $\Re_n$ 分解为 t 个($t\geqslant 1$，t 由 $\mathscr{T}$ 而定)不变子空间的直和，使每个子空间有一组基底如(23)之表达式，且这样的分解除同构者外是唯一的．

注意：定理 2 中所说的分解，一般是指复空间来说的，即在复空间内恒可这样分解，但在实空间内却不能保证必可作这样的分解，故对实空间，要保证能象这样分解，有时需要将它扩展为复空间，原因是初等因子要牵涉到特征值，即牵涉到求多项式之根的问题．

作为定理 2 之特例，又有：

定理3 矩阵 $\boldsymbol{A}$ 与对角矩阵相似的充要条件是：$\boldsymbol{A}$ 所对应的线性变换 $\mathscr{T}$ 可使全空间 $\Re_n$ 分解为 n 个一维不变子空间的直和．

[例 3] 讨论例 2 中的旋转变换．

解 因旋转变换 $\mathscr{T}$ 关于基底 $\boldsymbol{i}$，$\boldsymbol{j}$ 对应之矩阵为

$$\boldsymbol{A}=\begin{pmatrix}\cos\theta & \sin\theta\\ -\sin\theta & \cos\theta\end{pmatrix},$$

$\boldsymbol{A}$ 的两个特征值为 $\lambda_1=\cos\theta+i\sin\theta$，$\lambda_2=\cos\theta-i\sin\theta$，它们互异，故 $\boldsymbol{A}$ 与对角矩阵相似．因之，$\Re_2$ 可分解为两个一维不变子空间的直和，即平面中有两个向量 $\boldsymbol{e}_1$，$\boldsymbol{e}_2$，使[†]

$$\boldsymbol{e}_1\mathscr{T}=\lambda_1\boldsymbol{e}_1,\quad \boldsymbol{e}_2\mathscr{T}=\lambda_2\boldsymbol{e}_2.$$

[†] 读者可能对此有疑问，即平面中每个向量经旋转变换 $\mathscr{T}$ 后都应该变为与该向量不共线的向量(当 θ 不为 0° 或 180° 时)，而 $\boldsymbol{e}_1\mathscr{T}=\lambda_1\boldsymbol{e}_1$ 又确表示向量 $\boldsymbol{e}_1$ 经旋转 $\mathscr{T}$ 后与 $\boldsymbol{e}_1$ 共线．这又作何解释呢？这是因为这里已经把它扩展为复空间了．

第六章　二 次 齐 式

二次齐式的理论，起源于解析几何里二次曲线和二次曲面的研究. 在平面解析几何里，中心在坐标原点的二次有心曲线的方程是

$$Ax^2+2Bxy+Cy^2=F. \tag{1}$$

又，将坐标轴经旋转变换

$$\begin{cases} x=x'\cos\theta-y'\sin\theta, \\ y=x'\sin\theta+y'\cos\theta, \end{cases} \tag{2}$$

可使(1)式变为

$$A'x'^2+C'y'^2=F. \tag{3}$$

因为(1)、(3)两式中的常数项 F 是一样的，所以变换(2)只是将二次齐式 $Ax^2+2Bxy+Cy^2$ 变成 $A'x'^2+C'y'^2$，使乘积 $x'y'$ 的系数等于零. 显然，变换(2)是未知量的线性变换，也就是向量坐标的线性变换，这变换是满秩的，因为它的行列式等于

$$\begin{vmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{vmatrix}=1.$$

在许多应用方面，例如在物理学和力学中，都需要我们把二次齐式中未知量的个数扩充到 n 的情形，而且系数及未知量还不一定都取实数，即可以取复数. 本章讨论一般的二次齐式，讨论它如何经过满秩线性变换化为最简单形状，即所谓典型式，以及它的系数的变化规律.

第1节 化二次齐式为典型式

本节解决引言里所说的，怎样经过满秩线性变换，把 n 个变量的二次齐式变为 n 个变量的平方和的问题. 象这样只有变量平方项的二次齐式，叫做**典型式**.

n 个变量 $x_1, \cdots, x_n$ 的二次齐式，一般写为

$$\varphi(x_1, x_2, \cdots, x_n) = \sum_{i,j=1}^{n} a_{ij}x_ix_j$$

之形式. 若令

$$\boldsymbol{X} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \quad \boldsymbol{A} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix},$$

则易验证：$\varphi(x_1, x_2, \cdots, x_n)$ 可表写为矩阵之乘积的形式，如

$$\varphi(x_1, x_2, \cdots, x_n) = \sum_{i,j=1}^{n} a_{ij}x_ix_j = \boldsymbol{X}'\boldsymbol{A}\boldsymbol{X}, \tag{4}$$

这里 $\boldsymbol{X}'$ 是 $\boldsymbol{X}$ 的转置矩阵，即

$$\boldsymbol{X}' = (x_1, x_2, \cdots, x_n).$$

注意：两个二次齐式虽是恒等式(即不论变量 $x_1, \cdots, x_n$ 取任意一组值，所得的函数值总是相等的)，却不一定各对应项之系数必相等，例如

$$x^2 + xy + 2yx + y^2 = (x, y)\begin{pmatrix} 1 & 1 \\ 2 & 1 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix},$$

$$x^2 + 2xy + yx + y^2 = (x, y)\begin{pmatrix} 1 & 2 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix},$$

而这两个齐式实际上都等于 $x^2+3xy+y^2$，故它们是恒等的，可是矩阵 $\begin{pmatrix}1&1\\2&1\end{pmatrix}\neq\begin{pmatrix}1&2\\1&1\end{pmatrix}$. 故从 $\boldsymbol{X}'\boldsymbol{A}\boldsymbol{X}\equiv\boldsymbol{X}'\boldsymbol{B}\boldsymbol{X}$，不能断定 $\boldsymbol{A}=\boldsymbol{B}$，即可能有 $\boldsymbol{A}\neq\boldsymbol{B}$. 但因 $x_ix_j=x_jx_i$，故为求简便，假设 $a_{ij}=a_{ji}$，即在 $\varphi(x_1, x_2, \cdots, x_n)$ 中同类项合并后，$x_ix_j(i\neq j)$ 的系数记为 $2a_{ij}$ 或 $2a_{ji}$. 于是矩阵 $\boldsymbol{A}$ 具有性质 $\boldsymbol{A}=\boldsymbol{A}'$，有这样性质的矩阵叫对称矩阵. 关于对称矩阵说来，有

引理 设 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 都是对称矩阵 $(\boldsymbol{A}'=\boldsymbol{A}, \boldsymbol{B}'=\boldsymbol{B})$. 若 $\boldsymbol{X}'\boldsymbol{A}\boldsymbol{X}\equiv\boldsymbol{X}'\boldsymbol{B}\boldsymbol{X}$ (即 $\boldsymbol{X}'\boldsymbol{A}\boldsymbol{X}$ 与 $\boldsymbol{X}'\boldsymbol{B}\boldsymbol{X}$ 恒等)，则必 $\boldsymbol{A}=\boldsymbol{B}$.

证明 因 $\sum\limits_{i,j=1}^{n}a_{ij}x_ix_j=\boldsymbol{X}'\boldsymbol{A}\boldsymbol{X}\equiv\boldsymbol{X}'\boldsymbol{B}\boldsymbol{X}=\sum\limits_{i,j=1}^{n}b_{ij}x_ix_j$，故若令 $x_1=1, x_2=\cdots=x_n=0$ 代入，得 $a_{11}=b_{11}$. 同理，$a_{22}=b_{22}, \cdots, a_{nn}=b_{nn}$. 其次，若令 $x_1=x_2=1, x_3=\cdots=x_n=0$ 代入，又得 $a_{11}+a_{22}+a_{12}+a_{21}=b_{11}+b_{22}+b_{12}+b_{21}$，故由已证得之结果 $a_{ii}=b_{ii}(i=1, 2, \cdots)$ 及假设 $a_{ij}=a_{ji}$ 与 $b_{ij}=b_{ji}(i\neq j)$ 代入，即有 $2a_{12}=2b_{12}$，因之 $a_{12}=b_{12}$. 同理可知 $a_{st}=b_{st}(s\neq t)$. 这证明了 $\boldsymbol{A}=\boldsymbol{B}$. 证完.

今后，所谓二次齐式，总约定它相应的矩阵恒为对称矩阵，以后不另一一指明. 又因二次齐式(4)完全由对称矩阵 $\boldsymbol{A}$ 决定，所以矩阵 $\boldsymbol{A}$ 和它的秩分别称为二次齐式 $\varphi(x_1, \cdots, x_n)=\boldsymbol{X}'\boldsymbol{A}\boldsymbol{X}$ 的矩阵和它的秩.

再回到(4)式用满秩线性变换

$$\begin{cases}x_1=q_{11}y_1+q_{12}y_2+\cdots+q_{1n}y_n,\\x_2=q_{21}y_1+q_{22}y_2+\cdots+q_{2n}y_n,\\\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\x_n=q_{n1}y_1+q_{n2}y_2+\cdots+q_{nn}y_n,\end{cases}$$

即 $\boldsymbol{X}=\boldsymbol{Q}\boldsymbol{Y}$ 代入(4)中，这里

$$Q=\begin{pmatrix} q_{11} & q_{12} & \cdots & q_{1n} \\ q_{21} & q_{22} & \cdots & q_{2n} \\ \vdots & \vdots & & \vdots \\ q_{n1} & q_{n2} & \cdots & q_{nn} \end{pmatrix}, \quad Y=\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$$

则得

$$\varphi(x_1, \cdots, x_n)=\boldsymbol{X'AX}=\boldsymbol{Y'Q'AQY}=\boldsymbol{Y'BY},$$

其中 $\boldsymbol{B}=\boldsymbol{Q'AQ}=(b_{ij})$ 也一定是对称矩阵(这是因为 $\boldsymbol{B'}=(\boldsymbol{Q'AQ})'=\boldsymbol{Q'A'Q}=\boldsymbol{Q'AQ}=\boldsymbol{B}$),于是 $\boldsymbol{Y'BY}$ 是 n 个变量 y_1、y_2, $\cdots$, y_n 的二次齐式. 又,因为 $\boldsymbol{Q}$ 是满秩的,所以 $\boldsymbol{B}$ 的秩 $=\boldsymbol{A}$ 的秩. 故有:

定理 1 设有 n 个变量 $x_1, x_2, \cdots, x_n$ 的矩阵为 $\boldsymbol{A}$ 的二次齐式 $\boldsymbol{X'AX}$,对于变量 $\boldsymbol{X}=\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ 施行一个满秩矩阵 $\boldsymbol{Q}$ 之线性变换 $\boldsymbol{X}=\boldsymbol{QY}$ 后,得出一个 n 个新变量的二次齐式 $\boldsymbol{Y'BY}$,这里 $\boldsymbol{B}=\boldsymbol{Q'AQ}$,它的秩不变.

于是,(4)式经满秩线性变换 $\boldsymbol{X}=\boldsymbol{QY}$ 变为 $y_1, \cdots, y_n$ 的典型式 $b_1y_1^2+b_2y_2^2+\cdots+b_ny_n^2=\boldsymbol{Y'}\cdot\mathrm{diag}\{b_1, b_2, \cdots, b_n\}\cdot\boldsymbol{Y}$ 这问题,由定理 1 就意味着

$$\boldsymbol{Y'}\cdot\boldsymbol{Q'AQ}\cdot\boldsymbol{Y}\equiv\boldsymbol{Y'}\cdot\mathrm{diag}\{b_1, b_2, \cdots, b_n\}\cdot\boldsymbol{Y},$$

因为 $\boldsymbol{Q'AQ}$ 与 $\mathrm{diag}\{b_1, \cdots, b_n\}$ 都是对称的,根据引理,即得

$$\boldsymbol{Q'AQ}=\mathrm{diag}\{b_1, b_2, \cdots, b_n\}.$$

这就说明了:把二次齐式(4)化为典型式的问题,就等于求一个满秩矩阵 $\boldsymbol{Q}$,使 $\boldsymbol{Q'AQ}$($\boldsymbol{A}$ 是对称的)成为对角矩阵.

下面我们讨论这一问题,且一并来谈它的具体求法.

首先证明:一定存在关于变量 $x_1, \cdots, x_n$ 间的满秩变换,使二次齐式 $\varphi(x_1, \cdots, x_n)=\sum a_{ij}x_ix_j=\boldsymbol{X'AX}$ 变为新变量的二次齐式 $\boldsymbol{Y'BY}$ 中必至少有一个平方项出现. 这是因为:若

$\boldsymbol{X}'\boldsymbol{A}\boldsymbol{X}$ 中已有平方项，如 $a_{ii}\neq0$，问题就解决了；若不然，即当 $a_{11}=a_{22}=\cdots=a_{nn}=0$ 时，则至少有某一 $a_{ij}\neq0(i\neq j)$，例如 $a_{12}\neq0$，那末满秩变换

$$x_1=z_1+z_2,\quad x_2=z_1-z_2,\quad x_i=z_i(i=3,\ \cdots,\ n)$$

即
$$\begin{pmatrix}x_1\\x_2\\x_3\\\vdots\\x_n\end{pmatrix}=\begin{pmatrix}1&1&0&\cdots&0\\1&-1&0&\cdots&0\\0&0&1&\cdots&0\\\vdots&\vdots&\vdots&&\vdots\\0&0&0&\cdots&1\end{pmatrix}\begin{pmatrix}z_1\\z_2\\z_3\\\vdots\\z_n\end{pmatrix},$$

使二次齐式 $\boldsymbol{X}'\boldsymbol{A}\boldsymbol{X}$ 的项 $2a_{12}x_1x_2$ 变成了

$$2a_{12}x_1x_2=2a_{12}z_1^2-2a_{12}z_2^2,$$

其中出现了两个新变量 z_1 和 z_2 的平方项，它们的系数 $\neq0$，而且它们又不能够被其他的任何项消去（因在其余的每项中至少含有变量 $z_3,\ \cdots,z_n$ 中的一个）.

于是，只需考虑二次齐式 $\boldsymbol{X}'\boldsymbol{A}\boldsymbol{X}$ 中的平方项的系数 a_{ii} $(i=1,\ \cdots,\ n)$ 不都是 0 的情形．不损普遍性，可令 $a_{11}\neq0$. 因而将 $\boldsymbol{X}'\boldsymbol{A}\boldsymbol{X}$ 中所有含 x_1 的项提出，得

$$\boldsymbol{X}'\boldsymbol{A}\boldsymbol{X}=a_{11}x_1^2+2a_{12}x_1x_2+\cdots+2a_{1n}x_1x_n+\sum_{i,j=2}^{n}a_{ij}x_ix_j,$$

由于 $\varphi(x_1,\ \cdots,\ x_n)=\boldsymbol{X}'\boldsymbol{A}\boldsymbol{X}$ 中所有含 x_1 的项都全部包含在 $a_{11}^{-1}(a_{11}x_1+a_{12}x_2+\cdots+a_{1n}x_n)^2$ 之中，因此差

$$\varphi(x_1,\ x_2,\ \cdots,\ x_n)-a_{11}^{-1}(a_{11}x_1+a_{12}x_2+\cdots+a_{1n}x_n)^2$$

是一个只含变量 $x_2,\ \cdots,\ x_n$ 的二次齐式，表以 $\psi(x_2,\ \cdots,\ x_n)$，即 $\varphi(x_1,\ x_2,\ \cdots,\ x_n)=a_{11}^{-1}(a_{11}x_1+a_{12}x_2+\cdots+a_{1n}x_n)^2+\psi(x_2,\ \cdots,\ x_n)$．再作线性变换

$$y_1=a_{11}x_1+a_{12}x_2+\cdots+a_{1n}x_n,\quad y_i=x_i(i=2,\ \cdots,\ n),$$

即
$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \boldsymbol{P}_1 \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix},$$
因为 $\boldsymbol{P}_1$ 之行列式 $=a_{11}\neq 0$，故 $\boldsymbol{P}_1$ 是满秩的．于是，满秩变换 $\boldsymbol{X}=\boldsymbol{Q}_1\boldsymbol{Y}(\boldsymbol{Q}_1=\boldsymbol{P}_1^{-1})$ 化 $\varphi(x_1, \cdots, x_n)$ 为
$$\varphi(x_1, x_2, \cdots, x_n)=a_{11}^{-1}y_1^2+\psi(y_2, \cdots, y_n).$$
再，关于变量之个数 n 用归纳法，假定 $n-1$ 个变量 $y_2, \cdots, y_n$ 的二次齐式 $\psi(y_2, \cdots, y_n)$ 经过满秩变换：
$$\begin{pmatrix} y_2 \\ \vdots \\ y_n \end{pmatrix} = \boldsymbol{Q}_2\cdot\begin{pmatrix} z_2 \\ \vdots \\ z_n \end{pmatrix},\ (\text{其中 } \boldsymbol{Q}_2 \text{ 是 } n-1 \text{ 级满秩矩阵})$$
能够变成 $b_2z_2^2+\cdots+b_nz_n^2$（这里 $b_2, \cdots, b_n$ 可能有些是 0，也可能全都是零），因此满秩变换
$$\boldsymbol{Y}=\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} 1 & 0 \cdots 0 \\ 0 & \\ \vdots & \boldsymbol{Q}_2 \\ 0 & \end{pmatrix} \begin{pmatrix} z_1 \\ z_2 \\ \vdots \\ z_n \end{pmatrix}$$
与 $\boldsymbol{X}=\boldsymbol{Q}_1\boldsymbol{Y}$ 合并而得满秩变换 $\boldsymbol{X}=\boldsymbol{Q}\boldsymbol{Z}$，其中
$$\boldsymbol{Q}=\boldsymbol{Q}_1\begin{pmatrix} 1 & 0 \cdots 0 \\ 0 & \\ \vdots & \boldsymbol{Q}_2 \\ 0 & \end{pmatrix},\quad \boldsymbol{Z}=\begin{pmatrix} z_1 \\ z_2 \\ \vdots \\ z_n \end{pmatrix},$$
使 $\varphi(x_1, x_2, \cdots, x_n)=a_{11}^{-1}z_1^2+b_2z_2^2+\cdots+b_nz_n^2$．这就用归纳法证得了：$n$ 个变量 $x_1, \cdots, x_n$ 的二次齐式 $\boldsymbol{X}'\boldsymbol{A}\boldsymbol{X}$ 可经过满秩线性变换 $\boldsymbol{X}=\boldsymbol{Q}\boldsymbol{Y}$，变为新变量 $y_1, \cdots, y_n$ 的典型式；或相应地，对任意对称矩阵 $\boldsymbol{A}$，能找出一满秩矩阵 $\boldsymbol{Q}$，使 $\boldsymbol{Q}'\boldsymbol{A}\boldsymbol{Q}$ 为对角矩阵。

再来谈二次齐式化为典型式的几何意义.

回顾(1)式中 $Ax^2+2Bxy+Cy^2$ 的 x, y, 是动点 P 对于旧坐标系的坐标, 如图 6-1, 亦即向量 $\overrightarrow{OP}=x\boldsymbol{i}+y\boldsymbol{j}$, 但对新坐标系说来, 就有 $\overrightarrow{OP}=x'\boldsymbol{i}'+y'\boldsymbol{j}'$. 变换(2)表示同一向量 $\overrightarrow{OP}$ 在新旧坐标系下的坐标的变换.

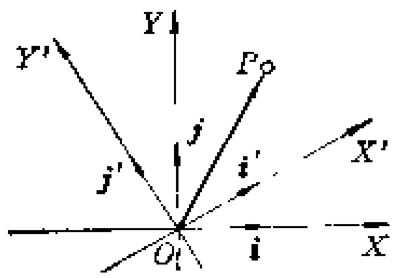


图 6-1

推广这个概念, 设有 n 维空间 $\mathfrak{R}_n$, 取它的一组基底 $\boldsymbol{e}_1, \cdots, \boldsymbol{e}_n$; 把 $x_1, \cdots, x_n$ 的二次齐式

$$\varphi(x_1, \cdots, x_n)=\sum_{i,j=1}^{n} a_{ij}x_ix_j=\boldsymbol{X}'\boldsymbol{A}\boldsymbol{X}$$

称为在 $\mathfrak{R}_n$ 内关于基底 $\boldsymbol{e}_1, \cdots, \boldsymbol{e}_n$ 的二次齐式, 其中 $x_1, \cdots, x_n$ 为 $\mathfrak{R}_n$ 中向量 $\boldsymbol{x}$ 对于基底 $\boldsymbol{e}_1, \cdots, \boldsymbol{e}_n$ 的坐标, 即 $\boldsymbol{x}=x_1\boldsymbol{e}_1+\cdots+x_n\boldsymbol{e}_n$.

把 $\varphi(x_1, \cdots, x_n)$ 化为 n 个变量 $y_1, \cdots, y_n$ 的平方和之意义, 就是要在仿射空间 $\mathfrak{R}_n$ 内找另一组基底 $\boldsymbol{f}_1, \cdots, \boldsymbol{f}_n$, 使二次齐式 $\varphi(x_1, \cdots, x_n)$ 对于每个向量

$$\boldsymbol{x}=x_1\boldsymbol{e}_1+\cdots+x_n\boldsymbol{e}_n=y_1\boldsymbol{f}_1+\cdots+y_n\boldsymbol{f}_n$$

的值, 可由公式

$$\varphi(x_1, x_2, \cdots, x_n)=b_1y_1^2+b_2y_2^2+\cdots+b_ny_n^2$$

来计算, 其中 $b_1, b_2, \cdots, b_n$ 是常数. 这就是说, 二次齐式 $\varphi(x_1, \cdots, x_n)$ 对于仿射空间 $\mathfrak{R}_n$ 内的基底 $\boldsymbol{f}_1, \cdots, \boldsymbol{f}_n$ 变成了典型式. 象这样的基底 $\boldsymbol{f}_1, \cdots, \boldsymbol{f}_n$ 叫做齐式 $\varphi(x_1, \cdots, x_n)$ 的典型基底, 而 $b_1, b_2, \cdots, b_n$ 叫做齐式 $\varphi(x_1, \cdots, x_n)$ 的典型系数.

因为已证得了必有满秩线性变换 $\boldsymbol{X}=\boldsymbol{Q}\boldsymbol{Y}$, 使

$$\boldsymbol{X}'\boldsymbol{A}\boldsymbol{X}=b_1y_1^2+\cdots+b_1y_n^2,$$

故如从关系式 $(\boldsymbol{f}_1, \cdots, \boldsymbol{f}_n)=(\boldsymbol{e}_1, \cdots, \boldsymbol{e}_n)\cdot\boldsymbol{Q}$ 而作 n 个向量 $\boldsymbol{f}_i(i=1, 2, \cdots, n)$，则由 $\boldsymbol{Q}$ 之满秩性即知 $\boldsymbol{f}_i(i=1, \cdots, n)$ 为 $\mathfrak{R}_n$ 的一组基底，于是若令向量 $\boldsymbol{x}=(\boldsymbol{e}_1, \cdots, \boldsymbol{e}_n)\cdot\boldsymbol{X}$ 关于基底 $\boldsymbol{f}_i(i=1, \cdots, n)$ 之坐标为 $\boldsymbol{Z}=\begin{pmatrix}z_1\\ \vdots\\ z_n\end{pmatrix}$，即

$$\boldsymbol{x}=(\boldsymbol{e}_1, \cdots, \boldsymbol{e}_n)\cdot\boldsymbol{X}=(\boldsymbol{f}_1, \cdots, \boldsymbol{f}_n)\cdot\boldsymbol{Z},$$

那末 $\boldsymbol{x}=(\boldsymbol{e}_1, \cdots, \boldsymbol{e}_n)\cdot\boldsymbol{X}=(\boldsymbol{f}_1, \cdots, \boldsymbol{f}_n)\cdot\boldsymbol{Z}=(\boldsymbol{e}_1, \cdots, \boldsymbol{e}_n)\cdot\boldsymbol{QZ}$，故必 $\boldsymbol{X}=\boldsymbol{QZ}$，因之 $\boldsymbol{QZ}=\boldsymbol{QY}$，再由 $\boldsymbol{Q}$ 之满秩性，得 $\boldsymbol{Z}=\boldsymbol{Y}$，说明了 $\boldsymbol{x}=y_1\boldsymbol{f}_1+\cdots+y_n\boldsymbol{f}_n$，即 $\boldsymbol{f}_1, \cdots, \boldsymbol{f}_n$ 为典型基底.

总括上述，证得了下面的三个等价性质：

定理 2（基本定理） i) *n 维仿射空间 $\mathfrak{R}_n$ 内的二次齐式 $\varphi(x_1, \cdots, x_n)$ 有典型基底；*

ii) n 个变量 $x_1, \cdots, x_n$ 的二次齐式 $\varphi(x_1, \cdots, x_n)$ 经过满秩线性变换 $\boldsymbol{X}=\boldsymbol{QY}$ 可变为新变量 $y_1, \cdots, y_n$ 的典型式；

iii) 任何对称矩阵 $\boldsymbol{A}$，能至少找到一个满秩矩阵 $\boldsymbol{Q}$，使 $\boldsymbol{Q'AQ}$ 为对角矩阵。

显然，基本定理中，对角矩阵 $\boldsymbol{Q'AQ}$ 的对角线上不为 0 的元素之个数等于矩阵 $\boldsymbol{A}$ 的秩. 当二次齐式 $\varphi(x_1, \cdots, x_n)$ 的系数 a_{ij} 都是实数的时候，可以限制变量 $x_1, \cdots, x_n$ 都取实数值，也就是说仿射空间 $\mathfrak{R}_n$ 是实空间，这时就叫 $\varphi(x_1, \cdots, x_n)$ 为实二次齐式. 由上面的证明易知，这时基本定理中满秩矩阵 $\boldsymbol{Q}$ 的元素都是实数. 如果 $\varphi(x_1, \cdots, x_n)$ 之系数 a_{ij} 以及变量 $x_1, \cdots, x_n$ 可以取复数，就叫 $\varphi(x_1, \cdots, x_n)$ 为复二次齐式，这时所讨论的仿射空间 $\mathfrak{R}_n$ 为复空间.

基本定理的证明方法，是顺次分出平方项，这也提供了把二次齐式化为典型式的具体方法.

[例 1]　化二次齐式 $\varphi(x_1, x_2, x_3)=2x_1x_2-6x_2x_3+2x_1x_3$ 为典型式.

解　因二次齐式 $\varphi(x_1, x_2, x_3)$ 缺变量的平方项，故首先施行满秩变换

$$x_1=y_1-y_2,\quad x_2=y_1+y_2,\quad x_3=y_3,$$

即
$$\begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix}=\begin{pmatrix}1&-1&0\\1&1&0\\0&0&1\end{pmatrix}\begin{pmatrix}y_1\\y_2\\y_3\end{pmatrix}=\boldsymbol{Q}_1\begin{pmatrix}y_1\\y_2\\y_3\end{pmatrix},$$

其中
$$\boldsymbol{Q}_1=\begin{pmatrix}1&-1&0\\1&1&0\\0&0&1\end{pmatrix},$$

使
$$\varphi(x_1, x_2, x_3)=(x_1, x_2, x_3)\cdot\boldsymbol{A}\cdot\begin{pmatrix}x_1\\x_2\\x_2\end{pmatrix},$$

其中
$$\boldsymbol{A}=\begin{pmatrix}0&1&1\\1&0&-3\\1&-3&0\end{pmatrix},$$

变成

$$\varphi(x_1, x_2, x_3)=2y_1^2-2y_2^2-4y_1y_3-8y_2y_3=(y_1, y_2, y_3)\boldsymbol{B}\begin{pmatrix}y_1\\y_2\\y_3\end{pmatrix},$$

其中
$$\boldsymbol{B}=\begin{pmatrix}2&0&-2\\0&-2&-4\\-2&-4&0\end{pmatrix}=\boldsymbol{Q}_1'\boldsymbol{A}\boldsymbol{Q}_1.$$

因 y_1^2 之系数不为 0，故施行满秩变换

$$z_1=2y_1-2y_3,\quad z_2=y_2,\quad z_3=y_3,$$

即 $$\begin{pmatrix}y_1\\y_2\\y_3\end{pmatrix}=\begin{pmatrix}1/2&0&1\\0&1&0\\0&0&1\end{pmatrix}\begin{pmatrix}z_1\\z_2\\z_3\end{pmatrix}=\boldsymbol{Q}_2\cdot\begin{pmatrix}z_1\\z_2\\z_3\end{pmatrix},$$

其中 $$\boldsymbol{Q}_2=\begin{pmatrix}1/2&0&1\\0&1&0\\0&0&1\end{pmatrix},$$

使 $$\varphi(x_1,\ x_2,\ x_3)=(y_1,\ y_2,\ y_3)\cdot\boldsymbol{B}\cdot\begin{pmatrix}y_1\\y_2\\y_3\end{pmatrix}$$

变成

$$\varphi(x_1,\ x_2,\ x_3)=\frac{1}{2}z_1^2-2z_2^2-2z_3^2-8z_2z_3$$

$$=(z_1,\ z_2,\ z_3)\cdot\boldsymbol{C}\cdot\begin{pmatrix}z_1\\z_2\\z_3\end{pmatrix},$$

但 $$\boldsymbol{C}=\begin{pmatrix}1/2&0&0\\0&-2&-4\\0&-4&-2\end{pmatrix}=\boldsymbol{Q}_2'\boldsymbol{B}\boldsymbol{Q}_2=\boldsymbol{Q}_2'\boldsymbol{Q}_1'\boldsymbol{A}\boldsymbol{Q}_1\boldsymbol{Q}_2.$$

又因 $-2z_2^2-2z_3^2-8z_2z_3$ 中 z_2^2 之系数不为0，故又施行满秩变换

$$t_1=z_1,\quad t_2=-2z_2-4z_3,\quad t_3=z_3,$$

即 $$\begin{pmatrix}z_1\\z_2\\z_3\end{pmatrix}=\begin{pmatrix}1&0&0\\0&-1/2&-2\\0&0&1\end{pmatrix}\begin{pmatrix}t_1\\t_2\\t_3\end{pmatrix}=\boldsymbol{Q}_3\begin{pmatrix}t_1\\t_2\\t_3\end{pmatrix},$$

其中 $$\boldsymbol{Q}_3=\begin{pmatrix}1&0&0\\0&-1/2&-2\\0&0&1\end{pmatrix},$$

使 $$\varphi(x_1, x_2, x_3)=(z_1, z_2, z_3)\cdot C\cdot\begin{pmatrix}z_1\\z_2\\z_3\end{pmatrix}$$

变成 $$\varphi(x_1, x_2, x_3)=\frac{1}{2}t_1^2-\frac{1}{2}t_2^2+6t_3^2$$

$$=(t_1, t_2, t_3)\cdot D\cdot\begin{pmatrix}t_1\\t_2\\t_3\end{pmatrix},$$

其中 $$D=\begin{pmatrix}1/2 & 0 & 0\\0 & -1/2 & 0\\0 & 0 & 6\end{pmatrix},$$

此即所求的典型式，而

$$D=Q_3'CQ_3=Q_3'Q_2'Q_1'AQ_1Q_2Q_3,$$

故变换矩阵为

$$Q=Q_1Q_2Q_3=\begin{pmatrix}1/2 & 1/2 & 3\\1/2 & -1/2 & -1\\0 & 0 & 1\end{pmatrix}.$$

化二次齐式为典型式的理论，与解析几何学中二次有心曲线或曲面的理论相类似，但却不能作为后而这个理论的推广，原因就是这里允许用任何一个满秩线性变换，而在化二次曲线或曲面到典型式时所用的满秩线性变换却是特殊的一种满秩变换，即本章开头所说的(2)式的旋转变换．然面，这一个几何理论也确可推广到 n 个变量的实二次齐式上去，这推广叫做化二次齐式到主轴上去(或者叫做二次齐式的主轴问题，有时又简称主轴问题)，这将在下章(欧氏空间)中详细地加以讨论。

练习二十五

1. 试化二次齐式 $\varphi(x_1, x_2, x_3)=\sum_{i,j=1}^{3} a_{ij}x_ix_j$ 为典型式，这里 $a_{ij}=|i-j|$.

2. 设 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 是同级的对称矩阵. 试证 $\boldsymbol{AB}$ 为对称的充要条件是 $\boldsymbol{AB}=\boldsymbol{BA}$.

3. 矩阵 $\boldsymbol{A}=(a_{ij})$ 之元具性质 $a_{ij}=-a_{ji}$(即 $\boldsymbol{A}'=-\boldsymbol{A}$)时，叫 $\boldsymbol{A}$ 为反对称矩阵. 证明：任一 n 级矩阵 $\boldsymbol{A}$ 可表示为一对称矩阵与一反对称矩阵的和，且这样的表示是唯一的.

4. 实对称矩阵之秩为1时，其主对角线上凡不等于零之元必为同符号的数.

第2节 惯性定理

本节仅考虑实空间 $\mathfrak{R}(=\mathfrak{R}_n)$. 由第1节基本定理之证明，我们会发现二次齐式的典型式不是唯一地确定的，换言之，它的典型基底或典型系数不止一组，这是因为我们可有多种方法化为平方和. 例如，二次齐式中平方项之系数有两个或多个不为零时，变换矩阵之选择就有许多. 现在要解决的问题是：典型系数虽不止一组，但每组中正、负的个数恒一定.

今设实二次齐式 $\varphi(x_1, x_2, \cdots, x_n)=\sum_{1}^{n} a_{ij}x_ix_j=\boldsymbol{X}'\boldsymbol{A}\boldsymbol{X}$ 中 $\boldsymbol{X}'=(x_1, x_2, \cdots, x_n)$ 为实空间 $\mathfrak{R}$ 之向量 $\boldsymbol{x}$ 对于基底 $\boldsymbol{e}_1, \cdots, \boldsymbol{e}_n$ 之坐标$\left(\boldsymbol{x}=\sum_{i=1}^{n} x_i\boldsymbol{e}_i\right)$. 这二次齐式有两个典型基底 $\boldsymbol{f}_1, \cdots, \boldsymbol{f}_n$ 与 $\boldsymbol{g}_1, \cdots, \boldsymbol{g}_n$. 用 $y_1, \cdots, y_n$ 表向量 $\boldsymbol{x}$ 对基底 $\boldsymbol{f}_1, \cdots, \boldsymbol{f}_n$ 之坐标，用 $z_1, \cdots, z_n$ 表 $\boldsymbol{x}$ 对基底 $\boldsymbol{g}_1, \cdots, \boldsymbol{g}_n$ 之坐标. 它们的对应坐标变换公式是

$$\boldsymbol{X}=\boldsymbol{Q}_1\cdot\boldsymbol{Y}, \quad \boldsymbol{X}=\boldsymbol{Q}_2\cdot\boldsymbol{Z},$$

这里 $\boldsymbol{Q}_1$ 与 $\boldsymbol{Q}_2$ 都是 n 级满秩矩阵. 如果二次齐式 $\varphi(x_1, \cdots, x_n)$ 对于基底 $\boldsymbol{f}_1, \cdots, \boldsymbol{f}_n$ 具有形状

$$\varphi=b_1y_1^2+\cdots+b_ky_k^2-b_{k+1}y_{k+1}^2-\cdots-b_ry_r^2, \tag{5}$$

而对于基底 $\boldsymbol{g}_1, \cdots, \boldsymbol{g}_n$ 具有形状

$$\varphi=c_1z_1^2+\cdots+c_pz_p^2-c_{p+1}z_{p+1}^2-\cdots-c_rz_r^2, \tag{6}$$

其中 r 为 φ 之秩, 即 $r_A=r$, 且每 b_i 及每 c_i 都假定是正数. 我们现在就是要证明 $k=p$.

用反证法, 设 $k\neq p$. 不失普遍性, 可令 $k<p$. 今考虑

$$y_1=0, \cdots, y_k=0, z_{p+1}=0, \cdots, z_n=0. \tag{7}$$

因每 y 及每 z 都是 $x_1, x_2, \cdots, x_n$ 的线性组合, 所以(7)式是含 n 个未知量 $x_1, x_2, \cdots, x_n$ 的 $n-p+k(<n)$ 个线性齐次方程组, 因之(7)必有一个非零的解 $(x_1, x_2, \cdots, x_n)$, 这说明了适合(7)之非零向量 $\boldsymbol{x}=x_1\boldsymbol{e}_1+\cdots+x_n\boldsymbol{e}_n$ 是存在的. 另方面, 把这非零解代入(5)、(6), 即得

$$\varphi=-b_{k+1}y_{k+1}^2-\cdots-b_ry_r^2=c_1z_1^2+\cdots+c_pz_p^2,$$

故 φ 之值一方面不能为正数而另方面在同时又不能为负数, 故得 φ 等于零. 再, 由于 $c_1>0, \cdots, c_p>0$, 必会产生 $z_1=0, z_2=0, \cdots, z_p=0$; 于是再与(7)合并又得到 $\boldsymbol{x}=z_1\boldsymbol{g}_1+z_2\boldsymbol{g}_2+\cdots+z_n\boldsymbol{g}_n=\boldsymbol{0}$ 为零向量, 这显然与上述 $\boldsymbol{x}$ 为非零向量之结论相矛盾. 故 $k<p$ 不成立. 同理, $k>p$ 也不行. 这就证得了 $k=p$. 即证得了:

定理 1 (二次齐式的惯性定理) *在一个实二次齐式 $\varphi(x_1, \cdots, x_n)=\sum a_{ij}x_ix_j$ 之典型式中, 正系数之个数与负系数之个数, 是与典型基底的选择没有关系的, 即为二次齐式对满秩线性变换的不变量.*

正因为它们是不变量, 所以我们叫这定理为惯性定理; 其中正系数之个数称为*正惯性指数*, 通常表为 P, 负系数之个数

叫负惯性指数，通常表为 N，它们与秩 r 间的关系为 $P+N=r$. 又叫 $s=P-N$ 为二次齐式 $\varphi(x_1, \cdots, x_n)$ 的正负号差(或符号差).

推论 对于实二次齐式 $\varphi(x_1, \cdots, x_n)=\sum a_{ij}x_ix_j$，必存在典型基底，使其典型式为

$$\varphi(x_1, \cdots, x_n)=z_1^2+\cdots+z_P^2-z_{P+1}^2-\cdots-z_{P+N}^2,$$

即典型式的系数都是 ±1.

证明 因 $\varphi(x_1, \cdots, x_n)$ 有典型基底 $\boldsymbol{f}_1, \cdots, \boldsymbol{f}_n$ 使 $\varphi(x_1, \cdots, x_n)$化为典型式

$$\varphi(x_1, \cdots, x_n)=b_1y_1^2+\cdots+b_Py_P^2-b_{P+1}y_{P+1}^2-\cdots-b_{P+N}y_{P+N}^2,$$

式中 $y_1, \cdots, y_n$ 为向量 $\boldsymbol{x}$ 对于基底 $\boldsymbol{f}_1, \cdots, \boldsymbol{f}_n$ 的坐标：$\boldsymbol{x}=\sum_{i=1}^{n}y_i\boldsymbol{f}_i$，而每 b_i 为正数. 再作满秩变换 $z_i=\sqrt{b_i}y_i(i=1, \cdots, P+N)$，$z_i=y_i(i=P+N+1, \cdots, n)$，即

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}=\boldsymbol{Q}\cdot\begin{pmatrix} z_1 \\ z_2 \\ \vdots \\ z_n \end{pmatrix},$$

其中

$$\boldsymbol{Q}=\begin{pmatrix} \frac{1}{\sqrt{b_1}} & & & & & \\ & \ddots & & & & \\ & & \frac{1}{\sqrt{b_{P+N}}} & & & \\ & & & 1 & & \\ & & & & \ddots & \\ & & & & & 1 \end{pmatrix},$$

则得 $\varphi(x_1, \cdots, x_n)=z_1^2+\cdots+z_P^2-z_{P+1}^2-\cdots-z_{P+N}^2$，这时适合

$(\boldsymbol{g}_1, \cdots, \boldsymbol{g}_n)=(\boldsymbol{f}_1, \cdots, \boldsymbol{f}_n)\cdot\boldsymbol{Q}$ 的 n 个向量 $\boldsymbol{g}_1, \cdots, \boldsymbol{g}_n$ 即为所求之典型基底，而 $z_1, \cdots, z_n$ 是向量 $\boldsymbol{x}$ 对这新基底的坐标：

$$\boldsymbol{x}=z_1\boldsymbol{g}_1+z_2\boldsymbol{g}_2+\cdots+z_n\boldsymbol{g}_n.$$

推论证完。

在解析几何里，研究二次曲线或曲面的一般情况后，紧接着是再研究退化的情形，即它的代数式可分解为一次因式之积的问题。相应地，实二次齐式 $\varphi(x_1, \cdots, x_n)=\sum a_{ij}x_ix_j$ 在什么情况下可分解为两个实一次因式之乘积呢？首先，设

$$\sum a_{ij}x_ix_j=(\alpha_1x_1+\alpha_2x_2+\cdots+\alpha_nx_n)(\beta_1x_1+\beta_2x_2+\cdots+\beta_nx_n). \tag{8}$$

分如下两种情况分别加以讨论：

i) 两个因式 $\alpha_1x_1+\alpha_2x_2+\cdots+\alpha_nx_n$ 与 $\beta_1x_1+\beta_2x_2+\cdots+\beta_nx_n$ 是线性相关的。

这时，有实数 ρ 使 $\beta_i=\rho\alpha_i$, $(i=1, \cdots, n)$. 于是，$\sum a_{ij}x_ix_j=\rho(\alpha_1x_1+\alpha_2x_2+\cdots+\alpha_nx_n)^2$. 因我们所讨论的是 $\boldsymbol{A}=(a_{ij})$ 的秩 $r\neq 0$, 所以 $\alpha_1, \cdots, \alpha_n$ 中至少有一个不为 0, 不损普遍性，令 $\alpha_1\neq 0$, 于是满秩变换

$$y_1=\alpha_1x_1+\alpha_2x_2+\cdots+\alpha_nx_n,$$
$$y_i=x_i \quad (i=2, \cdots, n)$$

化(8)式为典型式 $\sum a_{ij}x_ix_j=\rho y_1^2$, 即 $r=1$.

ii) 两个因式 $\alpha_1x_1+\alpha_2x_2+\cdots+\alpha_nx_n$ 与 $\beta_1x_1+\beta_2x_2+\cdots+\beta_nx_n$ 是线性无关的。

这时，不失一般性，可令

$$\begin{vmatrix}\alpha_1 & \alpha_2\\ \beta_1 & \beta_2\end{vmatrix}\neq 0,$$

于是满秩变换

$$\begin{cases} y_1 = \alpha_1 x_1 + \alpha_2 x_2 + \cdots + \alpha_n x_n, \\ y_2 = \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_n x_n, \\ y_3 = \qquad\qquad x_3, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ y_n = \qquad\qquad\qquad x_n \end{cases}$$

化(8)式为 $\sum a_{ij}x_ix_j = y_1y_2$. 再作满秩变换

$$y_1 = z_1 - z_2, \quad y_2 = z_1 + z_2, \quad y_i = z_i, \ (i = 3, \cdots, n),$$

就得到
$$\sum a_{ij}x_ix_j = y_1y_2 = z_1^2 - z_2^2,$$
这证明了 $r = 2$, $s = P - N = 1 - 1 = 0$.

反之，先假定 $r = 1$. 这时二次齐式化为典型式应是 $\sum a_{ij}x_ix_j = \lambda y_1^2$ 的形状，这里 y_1 是 $x_1, x_2, \cdots, x_n$ 的线性组合：$y_1 = \alpha_1 x_1 + \cdots + \alpha_n x_n$，所以 $\sum a_{ij}x_ix_j = \lambda(\alpha_1 x_1 + \cdots + \alpha_n x_n)^2$. 这证明了 $\sum a_{ij}x_ix_j$ 可分解为两个线性相关的一次因式之积.

其次，假定 $r = 2$, $s = 0$. 这时，典型式为

$$\sum a_{ij}x_ix_j = y_1^2 - y_2^2 = (y_1 + y_2)(y_1 - y_2),$$

这里 y_1 与 y_2 都是 $x_1, \cdots, x_n$ 的线性组合：

$$y_1 = \lambda_1 x_1 + \cdots + \lambda_n x_n,$$
$$y_2 = \mu_1 x_1 + \cdots + \mu_n x_n,$$

且它们线性无关. 于是，

$$y_1 + y_2 = (\lambda_1 + \mu_1)x_1 + \cdots + (\lambda_n + \mu_n)x_n,$$
$$y_1 - y_2 = (\lambda_1 - \mu_1)x_1 + \cdots + (\lambda_n - \mu_n)x_n$$

也必线性无关（因从 $k(\lambda_i + \mu_i) + l(\lambda_i - \mu_i) = 0$ 得 $(k+l)\lambda_i + (k-l)\mu_i = 0$，从而有 $k + l = 0$, $k - l = 0$，故 $k = l = 0$）. 于是，这时

$$\sum a_{ij}x_ix_j = (y_1 + y_2)(y_1 - y_2)$$

表示着是两个线性无关的一次因式的乘积.

综上所述，即证得了：

定理2 非零的实二次齐式 $\sum a_{ij}x_ix_j$ 能分解为二个线性相关(无关)的一次因式乘积的充要条件是 $r=1(r=2, s=0)$，即齐式之秩等于1(秩等于2，符号差为零)。

练习二十六

1. 实二次齐式的秩 r 和正负号差 s 一定同为奇数或者同为偶数，并且 $-r \leqslant s \leqslant r$。

2. 证明：凡适合第1题之条件的任两个整数 r 和 $s(r>0)$ 可以当做一个非零实二次齐式的秩和正负号差。

3. 设实二次齐式 $\sum a_{ij}x_ix_j$ 的正、负惯性指数分别为 P 与 N。若 $\alpha_1, \alpha_2, \cdots, \alpha_P, \beta_1, \beta_2, \cdots, \beta_N$ 是任意 $P+N$ 个正数，试证：必存在典型基底 $\boldsymbol{f}_1, \cdots, \boldsymbol{f}_n$，使二次齐式的典型式为

$$\sum a_{ij}x_ix_j=\alpha_1y_1^2+\cdots+\alpha_Py_P^2-\beta_1y_{P+1}^2-\cdots-\beta_Ny_{P+N}^2.$$

4. 如果二次齐式 $\sum a_{ij}x_ix_j=\boldsymbol{X}'\boldsymbol{A}\boldsymbol{X}$ 是满秩的(即 $\boldsymbol{A}$ 为满秩)，叫二次齐式 $\sum \frac{A_{ij}}{a}x_ix_j=\boldsymbol{X}'\boldsymbol{A}^{-1}\boldsymbol{X}$($A_{ij}$ 为 a_{ij} 在行列式 $a=|\boldsymbol{A}|$ 内的代数余子式)为它的逆式。试证：一个满秩的实二次齐式和它的逆式有相同的符号差，因之它们的正、负惯性指数之个数分别相等。

5. 证明复空间内秩 r 的二次齐式 $\sum_1^n a_{ij}x_ix_j$ 可以化为典型式 $z_1^2+z_2^2+\cdots+z_r^2$，因之复二次齐式没有惯性定理，即任意 r 个非零的数 $\alpha_1, \cdots, \alpha_r$ 都可为一组典型系数。

6. 非零的复二次齐式 $\sum a_{ij}x_ix_j$ 能分解为两个线性相关或无关的一次因式乘积之充要条件是它的秩等于1，或等于2。

7. 设 $\boldsymbol{A}=\begin{pmatrix}1&2&3&4\\2&3&4&5\\3&4&5&6\\4&5&6&7\end{pmatrix}$，$\boldsymbol{X}=\begin{pmatrix}x_1\\x_2\\x_3\\x_4\end{pmatrix}$，问实二次齐式 $\boldsymbol{X}'\boldsymbol{A}\boldsymbol{X}$ 能否分解为一次因式之积？如能，试分解之，并利用所得结果而再由归纳法去考虑 $\boldsymbol{X}'\boldsymbol{A}\boldsymbol{X}$，式中

$$\boldsymbol{X}=\begin{pmatrix}x_1\\x_2\\\vdots\\x_n\end{pmatrix},\quad \boldsymbol{A}=\begin{pmatrix}1&2&3&\cdots&n\\2&3&4&\cdots&n+1\\3&4&5&\cdots&n+2\\\vdots&\vdots&\vdots&&\vdots\\n&n+1&n+2&\cdots&2n-1\end{pmatrix}.$$

8. 设 $L_i=C_{i1}x_1+\cdots+C_{in}x_n(i=1, 2, \cdots, p+q)$，每一 C_{ij} 都是实数，试证实二次齐式

$$\varphi(x_1, \cdots, x_n)=L_1^2+L_2^2+\cdots+L_p^2-L_{p+1}^2-\cdots-L_{p+q}^2$$

的正、负惯性指数分别不超过 p 与 q.

第3节 二次曲面、二次曲线的投影分类和仿射分类

这里，我们利用前节的理论来讨论二次曲面、二次曲线的投影分类和仿射分类.

1. 投影分类

我们知道，空间中二次曲面的方程可写成

$$\sum_{i,j=1}^{4} a_{ij}x_ix_j=\boldsymbol{X}'\boldsymbol{A}\boldsymbol{X}=0, \tag{9}$$

其中 $\boldsymbol{A}=(a_{ij})$ 为一个4级实对称矩阵，$\boldsymbol{X}'=(x_1, x_2, x_3, x_4)$ 表示点的投影坐标[†]. 由上节理论，知有满秩变换

$$\boldsymbol{X}=\boldsymbol{P}\boldsymbol{Y}, \quad \boldsymbol{Y}'=(y_1, y_2, y_3, y_4), \tag{10}$$

使二次齐式 $\boldsymbol{X}'\boldsymbol{A}\boldsymbol{X}$ 变为下列十四种典型式

$$\left.\begin{array}{r}x_1^2+x_2^2+x_3^2+x_4^2\\x_1^2+x_2^2+x_3^2-x_4^2\\x_1^2+x_2^2-x_3^2-x_4^2\\x_1^2-x_2^2-x_3^2-x_4^2\\-x_1^2-x_2^2-x_3^2-x_4^2\end{array}\right\}\quad\left.\begin{array}{r}x_1^2+x_2^2+x_3^2\\x_1^2+x_2^2-x_3^2\\x_1^2-x_2^2-x_3^2\\-x_1^2-x_2^2-x_3^2\end{array}\right\}\quad\left.\begin{array}{r}x_1^2+x_2^2\\x_1^2-x_2^2\\-x_1^2-x_2^2\end{array}\right\}\quad\left.\begin{array}{r}x_1^2\\-x_1^2\end{array}\right\}$$

---

† 假定 $\frac{x_1}{x_4}=x, \frac{x_2}{x_4}=y, \frac{x_3}{x_4}=z$，则叫不全为零的任四个数之一组 (x_1, x_2, x_3, x_4) 为空间中点 (x, y, z) 的投影坐标（当 $x_4\neq0$ 时）. 如果 $x_4=0$，叫 $(x_1, x_2, x_3, 0)$ 为无穷远点（但 x_1, x_2, x_3 不全为0）.

中的一种．但对方程(9)说来，因两个方程左端之各项仅相差符号时，可视为同一方程，而变换(10)能看做空间中的投影(直射)变换，故二次曲面(9)可经过投影变换变为下列八种曲面标准形

$x_1^2+x_2^2+x_3^2+x_4^2=0$……虚椭球面(虚曲面)

$x_1^2+x_2^2+x_3^2-x_4^2=0$……椭球面或双叶双曲面或椭圆抛物面

$x_1^2+x_2^2-x_3^2-x_4^2=0$……单叶双曲面或双曲抛物面

$x_1^2+x_2^2+x_3^2=0$……虚锥面

$x_1^2+x_2^2-x_3^2=0$……实锥面或柱面

$x_1^2+x_2^2=0$……一对虚平面

$x_1^2-x_2^2=0$……一对实平面

$x_1^2=0$……二个重合的平面

中的一种．

若所研究的为复空间的投影分类，则因在复空间中典型系数都可假定等于1，所以上述八种应合并一些面成下列四种：

$x_1^2+x_2^2+x_3^2+x_4^2=0$……常态二次曲面

$x_1^2+x_2^2+x_3^2=0$……二次锥面

$x_1^2+x_2^2=0$……一对不同的平面

$x_1^2=0$……重合平面

今将上面的结果用表格表示于下：

i）在实空间内二次曲面的投影分类表（r表示秩，s 表符号差）

r	s	标准方程	二次曲面的类型
4	4	$x_1^2+x_2^2+x_3^2+x_4^2=0$	虚曲面（常态的）
4	2	$x_1^2+x_2^2+x_3^2-x_4^2=0$	常态卵形曲面（或非直纹面）
4	0	$x_1^2+x_2^2-x_3^2-x_4^2=0$	常态非卵形曲面（或直纹面）
3	3	$x_1^2+x_2^2+x_3^2=0$	虚锥面
3	1	$x_1^2+x_2^2-x_3^2=0$	锥面及柱面（或有一个顶点的锥面）†
2	2	$x_1^2+x_2^2=0$	一对虚平面
2	0	$x_1^2-x_2^2=0$	一对实平面
1	1	$x_1^2=0$	重合平面

† 顶点或是有限远点或是无限远点。

ii）在复空间内二次曲面的投影分类表（r 表示秩）

r	标准方程	曲面的类型
4	$x_1^2+x_2^2+x_3^2+x_4^2=0$	常态曲面
3	$x_1^2+x_2^2+x_3^2=0$	锥面（变态曲面）
2	$x_1^2+x_2^2=0$	一对不同的平面
1	$x_1^2=0$	重合平面

用同样的方法，可以得到平面上二次曲线的投影分类，今将分类的结果用如下表格列出，其中 (x_1, x_2, x_3) 是点 (x, y) 的投影坐标$\left(这里\ \dfrac{x_1}{x_3}=x,\ \dfrac{x_2}{x_3}=y\right)$.

iii) 实空间（平面）内二次曲线的投影分类表（r, s 的意义同前）

r	s	标　准　方　程	曲 线 的 类 型
3	3	$x_1^2+x_2^2+x_3^2=0$	虚曲线（常态的）
3	1	$x_1^2+x_2^2-x_3^2=0$	常态二次曲线（椭圆、双曲线、抛物线）
2	2	$x_1^2+x_2^2=0$	一对虚直线
2	0	$x_1^2-x_2^2=0$	一对实直线
1	1	$x_1^2=0$	重合直线

iv) 复空间内二次曲线的投影分类表（r 表示秩）

r	标　准　方　程	曲　线　的　类　型
3	$x_1^2+x_2^2+x_3^2=0$	常态曲线
2	$x_1^2+x_2^2=0$	一对不同的直线
1	$x_1^2=0$	重合的直线

2. 仿射分类

我们知道，仿射空间（斜坐标系）内一般二次曲面方程是

$$\sum_{i,j=1}^{3} a_{ij}x_ix_j+2a_{14}x_1+2a_{24}x_2+2a_{34}x_3+a_{44}=0, \qquad (11)$$

其中 x_1, x_2, x_3 是 $\mathfrak{R}_3$ 中向量 $\boldsymbol{x}$ 对基底 $\boldsymbol{e}_1$, $\boldsymbol{e}_2$, $\boldsymbol{e}_3$ 的坐标，即

$$\boldsymbol{x}=x_1\boldsymbol{e}_1+x_2\boldsymbol{e}_2+x_3\boldsymbol{e}_3,$$

$$\boldsymbol{A}=\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

是对称矩阵．因 $\mathfrak{R}_3$ 有典型基底 $\boldsymbol{f}_1$, $\boldsymbol{f}_2$, $\boldsymbol{f}_3$:

$$(\boldsymbol{f}_1, \boldsymbol{f}_2, \boldsymbol{f}_3)=(\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3)\cdot\boldsymbol{Q},$$

$\boldsymbol{Q}$ 是满秩的，于是变换

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \boldsymbol{Q} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}$$

使
$$\sum_{i,j=1}^{3} a_{ij} x_i x_j = (x_1, x_2, x_3) \boldsymbol{A} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

$$= (y_1, y_2, y_3) \cdot \boldsymbol{Q}'\boldsymbol{A}\boldsymbol{Q} \cdot \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}$$

$$= \delta_1 y_1^2 + \delta_2 y_2^2 + \delta_3 y_3^2,$$

故
$$\boldsymbol{Q}'\boldsymbol{A}\boldsymbol{Q} = \begin{pmatrix} \delta_1 & & \\ & \delta_2 & \\ & & \delta_3 \end{pmatrix},$$

这里 $\delta_i = \pm 1$ 或 0, 但 $\delta_1, \delta_2, \delta_3$ 不能都为零.

因我们是研究方程 (11), 所以将它的各项都变号没有影响, 于是有

① $r_A = 3$. 这时, 方程(11)化为下述两式之一:

i) $y_1^2 + y_2^2 + y_3^2 + 2b_{14}y_1 + 2b_{24}y_2 + 2b_{34}y_3 + b_{44} = 0$,

ii) $y_1^2 + y_2^2 - y_3^2 + 2b_{14}y_1 + 2b_{24}y_2 + 2b_{34}y_3 + b_{44} = 0$.

② $r_A = 2$. 这时, 方程(11)化为下述两式之一:

i) $y_1^2 + y_2^2 + 2b_{14}y_1 + 2b_{24}y_2 + 2b_{34}y_3 + b_{44} = 0$,

ii) $y_1^2 - y_2^2 + 2b_{14}y_1 + 2b_{24}y_2 + 2b_{34}y_3 + b_{44} = 0$.

③ $r_A = 1$. 这时, 方程(11)化为

$$y_1^2 + 2b_{14}y_1 + 2b_{24}y_2 + 2b_{34}y_3 + b_{44} = 0.$$

在第①种情况, 即 $r_A = 3$ 时, 易证方程分别可写为:

i)' $(y_1 + b_{14})^2 + (y_2 + b_{24})^2 + (y_3 + b_{34})^2 + b = 0$,

ii)' $(y_1 + b_{14})^2 + (y_2 + b_{24})^2 - (y_3 - b_{34})^2 + b = 0$

的形状，再经过坐标轴(斜的)的平移与伸缩，可分别变为

$z_1^2+z_2^2+z_3^2+1=0$, 即 i)′ 中的 $b>0$……虚椭球面

$z_1^2+z_2^2+z_3^2-1=0$, 即 i)′ 中的 $b<0$……椭球面

$z_1^2+z_2^2+z_3^2=0$, 即 i)′ 中的 $b=0$……仅有一实点的虚锥面

$z_1^2+z_2^2-z_3^2+1=0$, 即 ii)′ 中的 $b>0$……双叶双曲面

$z_1^2+z_2^2-z_3^2-1=0$, 即 ii)′ 中的 $b<0$……单叶双曲面

$z_1^2+z_2^2-z_3^2=0$, 即 ii)′ 中的 $b=0$……二次锥面

的形状，即曲面共六种类型.

在第②种情况，即 $r_A=2$ 时，易证方程分别可变为

i)′ $(y_1+b_{14})^2+(y_2+b_{24})^2+2b_{34}y_3+b=0$,

ii)′ $(y_1+b_{14})^2-(y_2-b_{24})^2+2b_{34}y_3+b=0$

形. 若 $b_{34}\neq 0$, $2b_{34}y_3+b=2b_{34}\left(y_3+\dfrac{b}{2b_{34}}\right)$, 故再经过坐标轴的平移与伸缩，又分别变为

$$z_1^2+z_2^2+z_3=0\text{……椭圆抛物面}$$

$$z_1^2-z_2^2+z_3=0\text{……双曲抛物面}$$

形，即曲面有二个类型.

但在 $b_{34}=0$ 时，可先平移坐标轴，再适当地伸缩使 i)′, ii)′ 分别变为

$z_1^2+z_2^2+1=0$, 即 i)′ 中 $b>0$……虚柱面

$z_1^2+z_2^2-1=0$, 即 i)′ 中 $b<0$……椭圆柱面

$z_1^2+z_2^2=0$, 即 i)′ 中 $b=0$……一直线(二个虚平面相交于一实直线)

$z_1^2-z_2^2+1=0$, 即 ii)′ 中 $b\neq 0$……双曲柱面

$z_1^2-z_2^2=0$, 即 ii)′ 中 $b=0$……相交的二平面

形，表示的曲面有五种类型.

最后，在第③种情况，即 $r_A=1$ 时，方程可写为

$$(y_1+b_{14})^2+2b_{24}y_2+2b_{34}y_3+b=0.$$

若 $b_{24}=b_{34}=0$，则平移及适当伸缩坐标轴，使变为

$z_1^2+1=0(b>0)$……一对平行的虚平面

$z_1^2-1=0(b<0)$……一对平行的平面

$z_1^2=0(b=0)$………重合平面

共有三种类型.

若 b_{24} 与 b_{34} 中仅一为0（如 $b_{24}\neq 0$, $b_{34}=0$），则上方程可写为 $(y_1+b_{14})^2+2b_{24}\left(y_2+\dfrac{b}{2b_{24}}\right)=0$，再平移及适当伸缩坐标轴，就得

$$z_1^2+z_2=0\text{……抛物柱面}.$$

即这时只一个类型.

如果 $b_{24}b_{34}\neq 0$，则满秩变换

$$\begin{cases} y_1=y_1', \\ y_2=\dfrac{b_{24}}{2(b_{24}^2+b_{34}^2)}y_2'+2b_{34}y_3', \\ y_3=\dfrac{b_{34}}{2(b_{24}^2+b_{34}^2)}y_2'-2b_{24}y_3' \end{cases}$$

使方程

$$(y_1+b_{14})^2+2b_{24}y_2+2b_{34}y_3+b=0$$

变为

$$(y_1'+b_{14})^2+(y_2'+b)=0,$$

于是再平移坐标轴后，即得

$$z_1^2+z_2=0$$

形，它表示抛物柱面，上已论述过.

这样，总括得到实二次曲面的十七个仿射类：

v）实空间内二次曲面的仿射分类表

类别	标准方程	曲面的类型
I	$x_1^2+x_2^2+x_3^2+1=0$	虚椭球面
II	$x_1^2+x_2^2+x_3^2-1=0$	椭球面
III	$x_1^2+x_2^2-x_3^2+1=0$	双叶双曲面
IV	$x_1^2+x_2^2-x_3^2-1=0$	单叶双曲面
V	$x_1^2+x_2^2+x_3^2=0$	只有一实点的虚锥面
VI	$x_1^2+x_2^2-x_3^2=0$	二次锥面
VII	$x_1^2+x_2^2+x_3=0$	椭圆抛物面
VIII	$x_1^2-x_2^2+x_3=0$	双曲抛物面
IX	$x_1^2+x_2^2+1=0$	虚柱面
X	$x_1^2+x_2^2-1=0$	椭圆柱面
XI	$x_1^2-x_2^2+1=0$	双曲柱面
XII	$x_1^2+x_2=0$	抛物柱面
XIII	$x_1^2+x_2^2=0$	交一实直线的二个虚平面
XIV	$x_1^2-x_2^2=0$	相交之二平面
XV	$x_1^2+1=0$	一对平行的虚平面
XVI	$x_1^2-1=0$	一对平行的平面
XVII	$x_1^2=0$	重合平面

对于复空间之二次曲面，由于其二次齐式的典型系数可为任意数，则知上述的 I、II、III、IV 类可合并为一类 $x_1^2+x_2^2+x_3^2+1=0$; V、VI 两类合为一类 $x_1^2+x_2^2+x_3^2=0$; VII、VIII 又合为一类 $x_1^2+x_2^2+x_3=0$; IX、X、XI 得合并为 $x_1^2+x_2^2+1=0$; XIII、XIV 可合并为 $x_1^2+x_2^2=0$; XV、XVI 得合为 $x_1^2+1=0$. 因之得到二次曲面的八个复的仿射分类，列表于下.

vi) 复空间内二次曲面的仿射分类表

类别	标准方程	曲面的类型
I	$x_1^2+x_2^2+x_3^2+1=0$	有中心的二次曲面
II	$x_1^2+x_2^2+x_3^2=0$	二次锥面
III	$x_1^2+x_2^2+x_3=0$	抛物面
IV	$x_1^2+x_2^2+1=0$	有中心的柱面
V	$x_1^2+x_2=0$	抛物柱面
VI	$x_1^2+x_2^2=0$	相交的二平面
VII	$x_1^2+1=0$	平行的二平面
VIII	$x_1^2=0$	重合平面

我们可以用同样的方法讨论二次曲线的仿射分类，不管是在实空间内或在复空间内．现在只将分类的结果用表列出于下，至于它的详细推导留给读者去完成．

vii) 实空间内二次曲线的仿射分类表

类别	标准方程	曲线的类型
I	$x_1^2+x_2^2+1=0$	虚椭圆
II	$x_1^2+x_2^2-1=0$	椭圆
III	$x_1^2+x_2^2=0$	只有实交点的二个虚直线
IV	$x_1^2-x_2^2+1=0$	双曲线
V	$x_1^2-x_2^2=0$	相交二直线
VI	$x_1^2+x_2=0$	抛物线
VII	$x_1^2+1=0$	一对虚平行线
VIII	$x_1^2-1=0$	一对平行直线
IX	$x_1^2=0$	二重合直线

viii）复空间内二次曲线的仿射分类表

类　别	标　准　方　程	曲　线　的　类　型
I	$x_1^2+x_2^2+1=0$	有心二次曲线
II	$x_1^2+x_2=0$	无心二次曲线(抛物线)
III	$x_1^2+x_2^2=0$	相交之二直线
IV	$x_1^2+1=0$	一对平行直线
V	$x_1^2=0$	二个重合直线

练习二十七

1. 证明：在 n 维实空间内非零的二次齐式共有 $\frac{1}{2}n(n+3)$ 个类，但非零的二次齐式方程共有 $\frac{n(n+4)}{4}$ 个或 $\frac{n^2+4n-1}{4}$ 个类(由 n 是偶数或奇数而定)。

2. 在 n 维复空间内，非零的二次齐式及它们的方程类别怎样？

第4节　厄米特(Hermite)齐式

实二次齐式的理论中，有些结论对于一般的复二次齐式就没有意义．例如，一般的复二次齐式没有惯性定理，它的典型式由秩唯一地确定．但对于某类特别复二次齐式(叫做厄米特齐式)，却具有与实二次齐式同样的性质．厄米特齐式在应用上很重要，例如在量子力学的研究中就占有一定的地位．

n 维复空间 $\mathfrak{R}$ 内的厄米特齐式指的是 n 个变量 $x_1, \cdots, x_n$ 和它们的共轭变量 $\bar{x}_1, \cdots, \bar{x}_n$ 之间的下述二次齐式

$$\varphi(x_1, \cdots, x_n)=\sum a_{ij}\bar{x}_i x_j,$$

其中 $a_{ji}=\bar{a}_{ij}$．于是矩阵 $\boldsymbol{A}=(a_{ij})$ 具性质 $\overline{\boldsymbol{A}'}=\boldsymbol{A}$，这样的矩阵称为厄米特矩阵．因之，实对称矩阵为厄米特矩阵的一个特例，且由定义又知厄米特矩阵对角线上元素是实数，而实二次

齐式为厄米特齐式的特例. 厄米特齐式 $\sum a_{ij}\bar{x}_i x_j(a_{ij}=\bar{a}_{ji})$ 很显然可写为矩阵的表示式:

$$\sum a_{ij}\bar{x}_i x_j=(\bar{x}_1, \cdots, \bar{x}_n)\cdot \boldsymbol{A}\cdot\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}=\overline{\boldsymbol{X}}'\boldsymbol{A}\boldsymbol{X},$$

这里 $\boldsymbol{X}=\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$. 我们把 $x_1, \cdots, x_n$ 当作复空间 $\Re$ 内向量 $\boldsymbol{x}$ 对于基底 $\boldsymbol{e}_1, \cdots, \boldsymbol{e}_n$ 的坐标: $\boldsymbol{x}=x_1\boldsymbol{e}_1+\cdots+x_n\boldsymbol{e}_n$, 这样, 又叫 $\overline{\boldsymbol{X}}'\boldsymbol{A}\boldsymbol{X}$ 是 $\Re$ 内关于基底 $\boldsymbol{e}_1, \cdots, \boldsymbol{e}_n$ 的厄米特齐式. 厄米特矩阵 $\boldsymbol{A}$ 和它的秩分别叫做厄米特齐式的矩阵和它的秩.

完全同第 1、2 节一样, 可建立厄米特齐式的理论, 因之这里只作概括地叙述以及一些必要的简单证明.

首先, 所谓厄米特齐式 $\overline{\boldsymbol{X}}'\boldsymbol{A}\boldsymbol{X}$ 的**典型式**, 指的是要在空间 $\Re$ 内找一组基底 $\boldsymbol{f}_1, \cdots, \boldsymbol{f}_n$, 使厄米特齐式 $\varphi(x_1, \cdots, x_n)=\sum a_{ij}\bar{x}_i x_j$ 对于每向量 $\boldsymbol{x}=x_1\boldsymbol{e}_1+\cdots+x_n\boldsymbol{e}_n=y_1\boldsymbol{f}_1+\cdots+y_n\boldsymbol{f}_n$ 的值由公式

$$\varphi(x_1, \cdots, x_n)=b_1\bar{y}_1y_1+\cdots+b_n\bar{y}_ny_n$$

来计算, 其中 $b_1, \cdots, b_n$ 是与 $\boldsymbol{f}_1, \cdots, \boldsymbol{f}_n$ 有关的常数, 这样的基底 $\boldsymbol{f}_1, \cdots, \boldsymbol{f}_n$ 叫做厄米特齐式 $\varphi(x_1, \cdots, x_n)$ 的**典型基底**, 而 $b_1, \cdots, b_n$ 叫做这个齐式关于典型基底 $\boldsymbol{f}_1, \cdots, \boldsymbol{f}_n$ 的**典型系数**.

特别要指出: 不论变量 $x_1, \cdots, x_n$ 取何值, 厄米特齐式 $\overline{\boldsymbol{X}}'\boldsymbol{A}\boldsymbol{X}$ 总是实做, 厄米特矩阵 $\boldsymbol{A}$ 之行列式 $|\boldsymbol{A}|$ 也是实数[†].

---

† $\overline{|\boldsymbol{A}|}=|\overline{\boldsymbol{A}}|=|\boldsymbol{A}'|=|\boldsymbol{A}|\Rightarrow|\boldsymbol{A}|$ 为实数. 又因 $\overline{\boldsymbol{X}}'\boldsymbol{A}\boldsymbol{X}$ 是数, 故 $(\overline{\boldsymbol{X}}'\boldsymbol{A}\boldsymbol{X})'=\overline{\boldsymbol{X}}'\boldsymbol{A}\boldsymbol{X}$, 于是 $\overline{(\overline{\boldsymbol{X}}'\boldsymbol{A}\boldsymbol{X})}=\overline{(\overline{\boldsymbol{X}}'\boldsymbol{A}\boldsymbol{X})'}=(\boldsymbol{X}'\overline{\boldsymbol{A}}'\overline{\boldsymbol{X}})=\overline{\boldsymbol{X}}'\boldsymbol{A}\boldsymbol{X}$, 这就证明了 $\overline{\boldsymbol{X}}'\boldsymbol{A}\boldsymbol{X}$ 是实数.

同第1节中定理1的证明方法完全一样，可证：

定理1 设有 n 个变量 $x_1, \cdots, x_n$ 的厄米特齐式 $\overline{X}'AX$，对于它的变量施行一个满秩矩阵 Q 的线性变换

$$X=\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}=Q\cdot\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}=Q\cdot Y,$$

就得到一个有 n 个新变量 $y_1, \cdots, y_n$ 的厄米特齐式 $\overline{Y}'BY$，其矩阵 $B=\overline{Q}'AQ$. 因而秩未变.

注意：象第1节里引理中那样强的要求，对厄米特矩阵说来是没有必要的. 换言之，恒有：

引理 不论 A、B 是任何的矩阵，从 $\overline{X}'AX\equiv\overline{X}'BX$，恒得 $A=B$.

证明 若令 $A=(a_{ij})$，$B=(b_{ij})$，则从

$$\sum_{s,t=1}^{n} a_{st}\bar{x}_s x_t=\overline{X}'AX\equiv\overline{X}'BX=\sum_{s,t=1}^{n} b_{st}\bar{x}_s x_t$$

中令 $x_1=1$，$x_2=\cdots=x_n=0$ 代入之，即得 $a_{11}=b_{11}$，同理可知 $a_{22}=b_{22}$，$\cdots$，$a_{nn}=b_{nn}$. 其次，令 $x_1=x_2=1$，$x_3=\cdots=x_n=0$ 代入之，又得到 $a_{11}+a_{22}+a_{12}+a_{21}=b_{11}+b_{22}+b_{12}+b_{21}$，故利用 $a_{ii}=b_{ii}(i=1, \cdots, n)$ 之关系得到

$$a_{12}+a_{21}=b_{12}+b_{21}. \tag{12}$$

再用 $x_1=1$，$x_2=i=\sqrt{-1}$ 代入，则得

$$a_{11}+a_{22}+(a_{12}-a_{21})i=b_{11}+b_{22}+(b_{12}-b_{21})i,$$

故又有

$$a_{12}-a_{21}=b_{12}-b_{21}. \tag{13}$$

从(12)与(13)即得 $a_{12}=b_{12}$ 及 $a_{21}=b_{21}$.

同理可证 $a_{st}=b_{st}(s\neq t)$，这就证明了 $A=B$，引理证完.

于是，根据定理1，如有满秩变换 $X=QY$，使 $\overline{X}'AX=$

$\overline{\boldsymbol{Y}}'\overline{\boldsymbol{Q}}'\boldsymbol{AQ}\,\boldsymbol{Y}=b_1\bar{y}_1y_1+\cdots+b_n\bar{y}_ny_n$ 为典型式而有典型系数 b_1, $\cdots$, b_n 时，则因

$$b_1\bar{y}_1y_1+\cdots+b_n\bar{y}_ny_n=\overline{\boldsymbol{Y}}'\cdot\mathrm{diag}\{b_1,\ b_2,\ \cdots,\ b_n\}\cdot\boldsymbol{Y},$$

故 $\overline{\boldsymbol{Y}}'(\overline{\boldsymbol{Q}}'\boldsymbol{AQ})\boldsymbol{Y}=\overline{\boldsymbol{Y}}'\cdot\mathrm{diag}\{b_1,\ b_2,\ \cdots,\ b_n\}\cdot\boldsymbol{Y}$, 因而据引理则有 $\overline{\boldsymbol{Q}}'\boldsymbol{AQ}=\mathrm{diag}\{b_1,\ b_2,\ \cdots,\ b_n\}$. 故当 $\boldsymbol{A}$ 为厄米特矩阵时，$\overline{\boldsymbol{Q}}'\boldsymbol{AQ}$ 亦必这样，随而 $\mathrm{diag}\{b_1,\ b_2,\ \cdots,\ b_n\}$ 也是厄米特的，故 b_1, b_2, $\cdots$, b_n 为实数，即厄米特齐式之典型系数均为实数. 于是，求厄米特齐式 $\overline{\boldsymbol{X}}'\boldsymbol{AX}$ 之典型基底 $\boldsymbol{f}_1$, $\cdots$, $\boldsymbol{f}_n$ 之意义，就等于求满秩矩阵 $\boldsymbol{Q}$ 使得 $\overline{\boldsymbol{Q}}'\boldsymbol{AQ}=\boldsymbol{B}$ 为对角矩阵（因之为实的）.

现在采用第 1 节中类似的方法，具体地来找典型基底: 首先，如非零的厄米特齐式 $\phi(x_1,\ \cdots,\ x_n)=\sum a_{ij}\bar{x}_ix_j(a_{ij}=\bar{a}_{ji})$ 中 $a_{11}=a_{22}=\cdots=a_{nn}=0$, 则必有一 $a_{ij}\neq0(i\neq j)$, 例如 $a_{12}\neq0$, 易证满秩变换

$$x_1=a_{12}z_1-z_2,\qquad x_2=z_1+\bar{a}_{12}z_2,\qquad x_i=z_i(i=3,\ \cdots,\ n)$$

使 $\varphi(x_1,\ \cdots,\ x_n)$ 变为变量 z_1, z_2, $\cdots$, z_n 的厄米特齐式 $\overline{\boldsymbol{Z}}'\boldsymbol{BZ}$,

$\boldsymbol{Z}=\begin{pmatrix}z_1\\ \vdots\\ z_n\end{pmatrix}$, $\overline{\boldsymbol{B}}'=\boldsymbol{B}$, 其中 $b_{11}=-b_{22}=2a_{12}\bar{a}_{12}>0$.

这样，只需考虑厄米特齐式 $\sum a_{ij}\bar{x}_ix_j$ 中的 $a_{ii}(i=1,\ 2,\ \cdots,\ n)$ 不全为零的情况. 譬如说 $a_{11}\neq0$, 这时易证满秩变换

$$y_1=a_{11}x_1+a_{12}x_2+\cdots+a_{1n}x_n,\qquad y_i=x_i(i=2,\ \cdots,\ n)$$

使 $\qquad\varphi(x_1,\ \cdots,\ x_n)=a_{11}^{-1}\cdot\bar{y}_1y_1+\Psi(y_2,\ \cdots,\ y_n)$,

而 $\Psi(y_2,\ \cdots,\ y_n)$ 是 $n-1$ 个变量 y_2, $\cdots$, y_n 的厄米特齐式，于是对变量的个数 n 用归纳法，证得了下面几个等价的:

定理 2 （基本定理）i) n 维复空间 $\mathfrak{R}$ 内的厄米特齐式 $\sum a_{ij}\bar{x}_ix_j(a_{ij}=\bar{a}_{ji})$ 一定有典型基底;

ii) n 个变量 $x_1, \cdots, x_n$ 的厄米特齐式 $\sum a_{ij}\bar{x}_i x_j$ 可以用满秩变换化为新变量 $y_1, \cdots, y_n$ 的典型式，如 $\sum a_{ij}\bar{x}_i x_j = b_1\bar{y}_1 y_1 + \cdots + b_r\bar{y}_r y_r$，其中 r 是厄米特齐式的秩（$b_1, \cdots, b_r$ 一定是实数）；

iii) 任何厄米特矩阵 $\boldsymbol{A}$，可以至少找得一个满秩矩阵 $\boldsymbol{Q}$，使所得的新厄米特矩阵 $\overline{\boldsymbol{Q}}'\boldsymbol{A}\boldsymbol{Q}$ 是对角矩阵（因之为实对角矩阵）.

注意：若 a_{ij} 全是实数，则 ii) 中所说的满秩变换的系数以及 iii) 中所说的满秩矩阵 $\boldsymbol{Q}$ 之元素也都可以为实数.

再假定厄米特齐式 $\sum a_{ij}\bar{x}_i x_j$ 有两个典型式：

$$b_1\bar{y}_1 y_1 + \cdots + b_k\bar{y}_k y_k - b_{k+1}\bar{y}_{k+1} y_{k+1} - \cdots - b_r\bar{y}_r y_r$$

与
$$c_1\bar{z}_1 z_1 + \cdots + c_p\bar{z}_p z_p - c_{p+1}\bar{z}_{p+1} z_{p+1} - \cdots - c_r\bar{z}_r z_r,$$

这里 r 是 $\boldsymbol{A}=(a_{ij})$ 的秩，b_i 与 $c_i (i=1, \cdots, r)$ 都是正数，则与第 2 节定理 1 之证明完全一样，能证 $k=p$，因此得到：

定理 3 （厄米特齐式的惯性定理） 一个厄米特齐式 $\sum a_{ij}\bar{x}_i x_j (a_{ij}=\bar{a}_{ji})$ 的典型式里的正、负系数之个数，与典型基底之选择是没有关系的．也就是说，它们是厄米特齐式对于满秩线性变换的不变量.

对于厄米特齐式的正、负惯性指数和正负号差，完全与实二次齐式之情形一样给以定义.

由定理 3，即可得知：

推论 在复空间 $\mathfrak{R}$ 内，任何厄米特齐式 $\sum a_{ij}\bar{x}_i x_j (a_{ij}=\bar{a}_{ji})$ 必有典型基底，使它的典型式为

$$\bar{z}_1 z_1 + \cdots + \bar{z}_P z_P - \bar{z}_{P+1} z_{P+1} - \cdots - \bar{z}_{P+N} z_{P+N},$$

其中 P, N 分别为它的正、负惯性指数.

证明的方法与第 2 节定理 1 之推论的证法一样.

练习二十八

1. 化厄米特齐式$\sum_{1}^{3} a_{st}\bar{x}_s x_t$为典型式，并求它的变换矩阵，这里已知$a_{st}=|s-t|+(s-t)i$.

2. 设$\boldsymbol{A}$与$\boldsymbol{B}$是两个n级厄米特矩阵，试证$\boldsymbol{AB}$也为厄米特矩阵的充要条件是$\boldsymbol{AB}=\boldsymbol{BA}$.

3. 若矩阵$\boldsymbol{A}=(a_{ij})$之元具性质$a_{ij}=-\bar{a}_{ji}$时(即$\bar{\boldsymbol{A}}'=-\boldsymbol{A}$)，叫$\boldsymbol{A}$为反厄米特矩阵. 证明：任一个$n$级矩阵可唯一地写为一个厄米特矩阵与一个反厄米特矩阵之和. 又，反厄米矩阵对角线上元素为零或纯虚数.

4. 证明：$\boldsymbol{A}$与$i\boldsymbol{A}$中，若有一为厄米特（反厄米特）矩阵，则另一必为反厄米特(厄米特)矩阵. 又，证明：n级反厄米特矩阵之行列式为实数或为纯虚数，全由n是偶数或奇数而定. 又，反厄米特矩阵$\boldsymbol{A}$之k次幂$\boldsymbol{A}^k$或仍为反厄米特矩阵或是厄米特的，全由k是奇数或偶数而定.

5. 秩等于1的厄米特矩阵对角线上凡不是零的数，都是具有同符号的实数；秩等于1的反厄米特矩阵对角线上凡不是零的数，必都具有虚部同符号的纯虚数.

6. 证明：满秩厄米特矩阵$\boldsymbol{A}$之逆矩阵$\boldsymbol{A}^{-1}$也是厄米特的. 又，反厄米特矩阵也有这样的性质.

7. 若厄米特齐式$\bar{\boldsymbol{X}}'\boldsymbol{A}\boldsymbol{X}$是满秩的，叫厄米特齐式$\bar{\boldsymbol{X}}'\boldsymbol{A}^{-1}\boldsymbol{X}$为$\bar{\boldsymbol{X}}'\boldsymbol{A}\boldsymbol{X}$之逆式. 证明：一个满秩厄米特齐式和它的逆式有相同的正负号差，即有相同个数的正、负惯性指数.

8. 设厄米特齐式$\sum a_{ij}\bar{x}_i x_j$之正负惯性指数分别为P, N; $\alpha_1, \cdots, \alpha_P, \beta_1, \cdots, \beta_N$为任$P+N$个正数. 证明：$\sum a_{ij}\bar{x}_i x_j$可有典型式$\alpha_1\bar{y}_1 y_1+\cdots+\alpha_P\bar{y}_P y_P-\beta_1\bar{y}_{P+1}y_{P+1}-\cdots-\beta_N\bar{y}_{P+N}y_{P+N}$.

第5节　有定及不定齐式及其应用

前已指出，实二次齐式和厄米特齐式之典型系数中正、负的个数，由该齐式自身所决定，而与典型基底之选择无关. 本节特讨论正、负惯性指数之个数与齐式之值的变化有何影响. 因实二次齐式为厄米特齐式的特例，所以下面只讨论厄米特的.

设 $\sum_{i,j=1}^{n} a_{ij}\bar{x}_i x_j = \bar{X}'AX$ 为厄米特齐式(满秩或降秩). 当它是满秩 ($r=n$) 时, 正惯性指数 P 变化的情况有两个特殊的 (即 $P=n$ 与 $P=0$) 和一个一般的 (即 $0<P<r(=n)$), 当它是降秩($r<n$)时, P 之变化状况也是有两个特殊的 (即 $P=r$ 与 $P=0$) 和一个一般的(即 $0<P<r(<n)$). 所以, 可将厄米特齐式分为下面五类情况来研究:

① $P=r=n$; ② $P=0$, $r=n$; ③ $P=r<n$; ④ $P=0$, $r<n$; ⑤ $0<P<r(\leqslant n)$.

现在分别讨论于下.

① $P=r=n$.

这时, 厄米特齐式之典型式可写为

$$\bar{X}'AX = \bar{y}_1 y_1 + \bar{y}_2 y_2 + \cdots + \bar{y}_n y_n.$$

于是, 只要 $y_1, \cdots, y_n$ 不全为 0, 则 $\bar{X}'AX>0$; 只有在 $y_1=\cdots=y_n=0$ 时, $\bar{X}'AX=0$. 我们又知 $X=\begin{pmatrix}x_1\\ \vdots\\ x_n\end{pmatrix}=QY$, $Y=\begin{pmatrix}y_1\\ \vdots\\ y_n\end{pmatrix}$, Q 是满秩的, 故 $y_1=\cdots=y_n=0$ 之充要条件是 $x_1=\cdots=x_n=0$. 从而可知这时 $\sum a_{ij}\bar{x}_i x_j$ 之值 (恒为实数, 这在上面已说过) 恒为正数, 只有在 $x_1=\cdots=x_n=0$ 时其值才为零.

② $P=0$, $r=n$.

这时, 典型式为 $\sum a_{ij}\bar{x}_i x_j = -\bar{y}_1 y_1 - \bar{y}_2 y_2 - \cdots - \bar{y}_n y_n$. 与 ① 的讨论完全一样, 可知这时 $\sum a_{ij}\bar{x}_i x_j$ 之值恒为负数, 只有当所有的 x_i 都等于零时它的值才是零.

③ $P=r<n$.

这时, 典型式为

$$\sum a_{ij}\bar{x}_i x_j = \bar{y}_1 y_1 + \cdots + \bar{y}_r y_r,\ r<n,$$

故不论 $x_1, \cdots, x_n$ 为任何数, 恒有 $\sum a_{ij}\bar{x}_i x_j \geqslant 0$. 这时虽然 x_1,

$\cdots$, x_n 不全为零，仍有使 $\sum a_{ij}\bar{x}_i x_j=0$ 的可能．例如适合 $y_1=\cdots=y_r=0$，而 y_{r+1}，$\cdots$，y_n 不全为 0 的一组 x_1，$\cdots$，x_n 的值，就有 $\sum a_{ij}\bar{x}_i x_j=0$．注意与①之差别．

④ $P=0$，$r<n$．

这时，与③同样，由于典型式为

$$\sum a_{ij}\bar{x}_i x_j=-\bar{y}_1 y_1-\cdots-\bar{y}_r y_r,$$

则知不论 x_1，$\cdots$，x_n 是任何数，恒有 $\sum a_{ij}\bar{x}_i x_j\leqslant 0$．注意 x_1，$\cdots$，x_n 虽不全为零，也可使 $\sum a_{ij}\bar{x}_i x_j=0$，这是与②之差异．

⑤ $0<P<r(\leqslant n)$．

这时，典型式为

$$\sum a_{ij}\bar{x}_i x_j=\bar{y}_1 y_1+\cdots+\bar{y}_P y_P-\bar{y}_{P+1}y_{P+1}-\cdots-\bar{y}_r y_r,$$
$$(r=P+N).$$

于是适合 y_1，$\cdots$，y_P 不全为 0，而 $y_{P+1}=\cdots=y_n=0$ 的 x_1，$\cdots$，x_n 之值（当然这些 x_1，$\cdots$，x_n 也不全是零）就必使

$$\sum a_{ij}\bar{x}_i x_j>0;$$

又适合 y_{P+1}，$\cdots$，y_r 不全为 0，而 $y_1=\cdots=y_P=y_{r+1}=\cdots=y_n=0$ 的一组不全是 0 的 x_1，$\cdots$，x_n 的值就必使

$$\sum a_{ij}\bar{x}_i x_j<0.$$

也可能有不全为 0 之 x_1，$\cdots$，x_n 的一组值能使

$$\sum a_{ij}\bar{x}_i x_j=0,$$

例如 $r<n$ 时适合 y_{r+1}，$\cdots$，y_n 不全是 0 而 $y_1=\cdots=y_r=0$ 的一组不全为 0 的 x_1，$\cdots$，x_n 的值确使 $\sum a_{ij}\bar{x}_i x_j=0$．

归纳上述，得知厄米特齐式（实二次齐式）之值的变化状况有五种可能：

定义 1 设有厄米特齐式（实二次齐式）$\varphi(x_1, \cdots, x_n)=\sum a_{ij}\bar{x}_i x_j$，则：

i）若恒有 $\varphi(x_1, \cdots, x_n)>0$，且仅在 $x_1=\cdots=x_n=0$ 时才

有 $\varphi(x_1, \cdots, x_n)=0$, 我们叫 $\varphi(x_1, \cdots, x_n)$ 是正定的；

ii) 若恒有 $\varphi(x_1, \cdots, x_n)<0$, 且仅在 $x_1=\cdots=x_n=0$ 时才有 $\varphi(x_1, \cdots, x_n)=0$, 就叫 $\varphi(x_1, \cdots, x_n)$ 是负定的；

iii) 若恒有 $\varphi(x_1, \cdots, x_n)\geqslant 0$ (即 φ 之值能为零，虽然 $x_1, \cdots, x_n$ 不全为 0), 就叫 $\varphi(x_1, \cdots, x_n)$ 为半正定的；

iv) 若恒有 $\varphi(x_1, \cdots, x_n)\leqslant 0$ (即虽 $x_1, \cdots, x_n$ 不全是 0, 但 φ 之值能为零), 就叫 $\varphi(x_1, \cdots, x_n)$ 为半负定的；

v) 若 $\varphi(x_1, \cdots, x_n)$ 之值有时为正、有时又为负，就叫 $\varphi(x_1, \cdots, x_n)$ 是不定的.

象这样所分的类，叫做厄米特齐式的值类.

定义 2 厄米特矩阵(实对称矩阵) $\boldsymbol{A}$ 叫做正定、负定、半正定、半负定或不定，是由它所对应的齐式 $\overline{\boldsymbol{X}}'\boldsymbol{A}\boldsymbol{X}$ (或 $\boldsymbol{X}'\boldsymbol{A}\boldsymbol{X}$) 为正定、负定、半正定、半负定、或不定而来定义的.

据前面的讨论及上面的定义，易知：

定理 1 厄米特(实二次)齐式 $\sum\limits_{i,j=1}^{n} a_{ij}\bar{x}_i x_j$ 为

i) 正定的充要条件是 $P=r=n$;

ii) 负定的充要条件是 $P=0,\ r=n$;

iii) 半正定的充要条件是 $P=r<n$;

iv) 半负定的充要条件是 $P=0,\ r<n$;

v) 不定的充要条件是 $0<P<r(\leqslant n)$;

这里的 P 是齐式的正惯性指数，r 是齐式的秩.

下面再介绍用矩阵之主子式来判定厄米特齐式(实二次齐式)的值类的方法，叫做行列式判别法，在实际运用时常采用此法.

定理 2 厄米特(实二次)齐式 $\varphi(x_1, \cdots, x_n)=\sum a_{ij}\bar{x}_i x_j$ 为正定的充要条件是

$$a_{11}>0,\quad \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}>0,\quad \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}>0,$$

$$\cdots,\quad \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}>0. \tag{14}$$

先证明下列的:

引理 1 *若厄米特(实对称)矩阵 $\boldsymbol{A}$ 是正定的,则其行列式 $|\boldsymbol{A}|$ 为正数.*

证明 $\boldsymbol{A}$ 之正定性就说明有满秩矩阵 $\boldsymbol{Q}$, 使 $\overline{\boldsymbol{Q}}'\boldsymbol{A}\boldsymbol{Q}=\mathrm{diag}\{1, 1, \cdots, 1\}=\boldsymbol{E}$(定理 1 及前节定理 3 之推论),取其行列式,得 $1=|\overline{\boldsymbol{Q}}'|\cdot|\boldsymbol{A}|\cdot|\boldsymbol{Q}|=|\boldsymbol{A}|\cdot|\overline{\boldsymbol{Q}}|\cdot|\boldsymbol{Q}|$, 而 $|\overline{\boldsymbol{Q}}|\cdot|\boldsymbol{Q}|$ 是正数,故 $|\boldsymbol{A}|>0$. 引理 1 证完.

先证定理 2 中条件的必要性. 设 $\varphi(x_1, \cdots, x_n)=\sum a_{ij}\bar{x}_i x_j$ 是正定的,由引理 1 知有

$$|\boldsymbol{A}|=\begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix}>0.$$

又据正定之意义,知道 $\varphi(x_1, \cdots, x_{n-1}, 0)=\sum_{i,j=1}^{n-1} a_{ij}\bar{x}_i x_j$ 也是正定的,故再据引理 1 得到

$$\begin{vmatrix} a_{11} & \cdots & a_{1,n-1} \\ \vdots & & \vdots \\ a_{n-1,1} & \cdots & a_{n-1,n-1} \end{vmatrix}>0.$$

继续这种的方法,即陆续地利用 $\varphi(x_1, \cdots, x_{n-2}, 0, 0)$, $\cdots$, $\varphi(x_1, x_2, 0, \cdots, 0)$, $\varphi(x_1, 0, \cdots, 0)$ 之正定性,就能证明

$$\begin{vmatrix} a_{11} & \cdots & a_{1,n-2} \\ \vdots & & \vdots \\ a_{n-2,1} & \cdots & a_{n-2,n-2} \end{vmatrix} > 0, \ \cdots, \ \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} > 0, \ a_{11} > 0.$$

即证明了条件的必要性.

再证条件的充分性. 对于变量的个数,用归纳法. 当 $n=1$ 时,有 $\varphi(x_1)=a_{11}\bar{x}_1x_1$, 这时(14)式只含一个不等式 $a_{11}>0$, 于是当 $x_1\neq 0$ 时显见有 $\varphi(x_1)>0$, 即 $\varphi(x_1)$ 正定, 就证明了 $n=1$ 时的正确性. 再归纳地假定凡满足条件(14)的 $n-1$ 个变量的厄米特齐式一定是正定的, 而来考虑 (14) 式成立的 n 个变量 $x_1, \cdots, x_n$ 的厄米特齐式

$$\varphi(x_1, \cdots, x_n)=\sum_1^n a_{ij}\bar{x}_ix_j.$$

这时,因

$$\psi(x_1, \cdots, x_{n-1})=\varphi(x_1, \cdots, x_{n-1}, 0)=\sum_1^{n-1} a_{ij}\bar{x}_ix_j$$

是 $n-1$ 个变量 $x_1, \cdots, x_{n-1}$ 的厄米特齐式, 且有 (14) 中前 $n-1$ 个不等式, 故据归纳法的假设知道 $\psi(x_1, \cdots, x_{n-1})$ 是正定的, 于是可令其 $n-1$ 个典型系数都为 1, 即有 $n-1$ 级满秩矩阵 $\boldsymbol{Q}_1$ 使

$$\bar{\boldsymbol{Q}}_1'\cdot\begin{pmatrix} a_{11} & \cdots & a_{1,n-1} \\ \vdots & & \vdots \\ a_{n-1,1} & \cdots & a_{n-1,n-1} \end{pmatrix}\cdot\boldsymbol{Q}_1=\begin{pmatrix} 1 & & & \\ & 1 & & \\ & & \ddots & \\ & & & 1 \end{pmatrix},$$

于是满秩变换

$$\begin{pmatrix} x_1 \\ \vdots \\ x_{n-1} \\ x_n \end{pmatrix}=\boldsymbol{Q}_2\cdot\begin{pmatrix} y_1 \\ \vdots \\ y_{n-1} \\ y_n \end{pmatrix}, \text{ 其中 } \boldsymbol{Q}_2=\begin{pmatrix} \boldsymbol{Q}_1 & \boldsymbol{0} \\ \boldsymbol{0} & 1 \end{pmatrix},$$

使
$$\varphi(x_1, \cdots, x_n)=\bar{y}_1 y_1+\cdots+\bar{y}_{n-1} y_{n-1}$$
$$+(b_{1n}\bar{y}_1+\cdots+b_{n-1,n}\bar{y}_{n-1})y_n$$
$$+(b_{n1}y_1+\cdots+b_{n,n-1}y_{n-1})\bar{y}_n+b_{nn}\bar{y}_n y_n,$$

即
$$\bar{Q}_2' A Q_2=\begin{pmatrix} 1 & & & & b_{1n} \\ & 1 & & & b_{2n} \\ & & \ddots & & \vdots \\ & & & 1 & b_{n-1,n} \\ b_{n1} & b_{n2} & \cdots & b_{n,n-1} & b_{nn} \end{pmatrix}$$
为厄米特矩阵，因之 $b_{ij}=\bar{b}_{ji}$. 再作满秩变换
$$\begin{cases} z_1=y_1 & +b_{1n}y_n, \\ z_2=\quad y_2 & +b_{2n}y_n, \\ \cdots\cdots\cdots\cdots\cdots\cdots & \\ z_{n-1}= \quad y_{n-1} & +b_{n-1,n}y_n, \\ z_n= & y_n, \end{cases}$$
就得到
$$\varphi(x_1, \cdots, x_n)=(\bar{y}_1+b_{n1}\bar{y}_n)(y_1+b_{1n}y_n)+\cdots$$
$$+(\bar{y}_{n-1}+b_{n,n-1}\bar{y}_n)(y_{n-1}+b_{n-1,n}y_n)+c\bar{y}_n y_n$$
$$=\bar{z}_1 z_1+\cdots+\bar{z}_{n-1}z_{n-1}+c\bar{z}_n z_n,$$
其中 $c=b_{nn}-b_{1n}\bar{b}_{1n}-\cdots-b_{n-1,n}\bar{b}_{n-1,n}$ 为实数，这就是说，有满秩变换
$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}=Q\cdot\begin{pmatrix} z_1 \\ \vdots \\ z_n \end{pmatrix}$$
使
$$\bar{X}' A X=\sum a_{ij}\bar{x}_i x_j=(\bar{z}_1, \bar{z}_2, \cdots, \bar{z}_n)\cdot\bar{Q}' A Q\cdot\begin{pmatrix} z_1 \\ z_2 \\ \vdots \\ z_n \end{pmatrix}$$

$$=(\bar{z}_1, \bar{z}_2, \cdots, \bar{z}_n)\cdot\begin{pmatrix}1 & & & \\ & 1 & & \\ & & \ddots & \\ & & & 1 \\ & & & & C\end{pmatrix}\cdot\begin{pmatrix}z_1\\ z_2\\ \vdots\\ z_n\end{pmatrix},$$

即 $\overline{Q}'AQ=\mathrm{diag}\{1, \cdots, 1, C\}$，故取行列式得到

$$C=|A|\cdot|Q|\cdot|\overline{Q}|.$$

由(14)之最后一式的假设 $|A|>0$，就知道 $C>0$，这说明了厄米特齐式 $\varphi(x_1, \cdots, x_n)$ 的 n 个典型系数都是正数，故 $\varphi(x_1, \cdots, x_n)$ 是正定的，充分性获证。

于是定理 2 证完。

有了定理 2，不难获得负定的判别条件。这是因为 $\sum a_{ij}x_ix_j$ 为负定的充要条件显然是 $\sum(-a_{ij})\bar{x}_ix_j$ 为正定的，故由定理 2 可知这时之充要条件为

$$(-1)^k\cdot\begin{vmatrix}a_{11} & \cdots & a_{1k}\\ \vdots & & \vdots\\ a_{k1} & \cdots & a_{kk}\end{vmatrix}>0, \quad (k=1, 2, \cdots, n).$$

故又证得了：

定理 3 厄米特(实对称)矩阵 $A=(a_{ij})$ 为负定的充要条件是

$$a_{11}<0, \quad \begin{vmatrix}a_{11} & a_{12}\\ a_{21} & a_{22}\end{vmatrix}>0,$$

$$\begin{vmatrix}a_{11} & a_{12} & a_{13}\\ a_{21} & a_{22} & a_{23}\\ a_{31} & a_{32} & a_{33}\end{vmatrix}<0, \cdots, (-1)^n\begin{vmatrix}a_{11} & a_{12} & \cdots & a_{1n}\\ a_{21} & a_{22} & \cdots & a_{2n}\\ \vdots & \vdots & & \vdots\\ a_{n1} & a_{n2} & \cdots & a_{nn}\end{vmatrix}>0.$$

由定理 2 与 3，易知下面的：

推论 厄米特(实对称)矩阵为正定的充要条件是它的一

切主子式都是正数；它为负定的充要条件是一切偶数级主子式为正数，而一切奇数级主子式为负数。

应用上面的行列式之判别方法来决定厄米特(实二次)齐式所属的值类，有时是很方便的，例如有三个变量的二次齐式

$$\varphi=10x^2-2y^2+3z^2+4zx+4xy,$$

其矩阵是
$$A=\begin{pmatrix}10 & 2 & 2\\ 2 & -2 & 0\\ 2 & 0 & 3\end{pmatrix}.$$

因为
$$10>0,\quad \begin{vmatrix}10 & 2\\ 2 & -2\end{vmatrix}=-24<0,$$
$$\begin{vmatrix}10 & 2 & 2\\ 2 & -2 & 0\\ 2 & 0 & 3\end{vmatrix}=-64<0,$$

所以 φ 不是正定的也不是负定的；但又因它为满秩的，故又不能是半正定或半负定. 于是 φ 是不定的.

有了定理 2 与 3，则易知半定的判别方法为：

定理 4 厄米特(实对称)矩阵为半正定的充要条件是它为降秩的，且一切主子式都为非负的数.

定理 5 厄米特(实对称)矩阵为半负定的充要条件是它为降秩的，且一切偶级主子式为非负的数，一切奇级主子式为非正的数.

若能证明定理 4，则不难证明定理 5. 这是由于 $\sum a_{ij}\bar{x}_ix_j$ 为半负定的充要条件是 $\sum(-a_{ij})\bar{x}_ix_j$ 为半正定的原故. 于是我们今后只需证定理 4.

证明 设 $A=(a_{ij})$. 先证必要性. 若 A 为半正定，则由定理 1 知 A 是降秩，即 $r=r_A<n$，且又恒有

$$\varphi(x_1,\ \cdots,\ x_n)=\sum a_{ij}\bar{x}_ix_j=\bar{X}'AX\geqslant 0.$$

今取 $\boldsymbol{A}$ 之任一个 m 级的主子矩阵

$$\boldsymbol{A}^{(m)}=\begin{pmatrix} a_{k_1k_1} & \cdots & a_{k_1k_m} \\ \vdots & & \vdots \\ a_{k_mk_1} & \cdots & a_{k_mk_m} \end{pmatrix},$$

而 $\boldsymbol{A}^{(m)}$ 所对应的厄米特(实二次)齐式用

$$\psi(x_{k_1},\ \cdots,\ x_{k_m})=\sum_{s,t=1}^{m} a_{k_sk_t}\bar{x}_{k_s}x_{k_t}$$

来表示．显然，$\psi(x_{k_1},\ \cdots,\ x_{k_m})$ 之值就是在 $\varphi(x_1,\ \cdots,\ x_n)$ 中令所有那些 $i\neq k_1,\ \cdots,\ k_m$ 的 x_i 都是零而得到的，所以也恒有 $\psi(x_{k_1},\ \cdots,\ x_{k_m})\geqslant 0$，这是说 ψ 为正定的或半正定的，因而其典型系数都可以假定为 1 和 0（当它是半正定时，0 必然要出现）．于是，有 m 级满秩矩阵 $\boldsymbol{Q}$，使

$$\overline{\boldsymbol{Q}}'\boldsymbol{A}^{(m)}\boldsymbol{Q}=\mathrm{diag}\{\delta_1,\ \delta_2,\ \cdots,\ \delta_m\},$$

式中每 $\delta_i=1$ 或 0，再取两边之行列式，则得

$$|\boldsymbol{A}^{(m)}|\cdot|\boldsymbol{Q}|\cdot|\overline{\boldsymbol{Q}}|=\delta_1\delta_2\cdots\delta_m\geqslant 0,$$

故 $|\boldsymbol{A}^{(m)}|\geqslant 0$（因 $|\boldsymbol{Q}|\cdot|\overline{\boldsymbol{Q}}|>0$），这就证明了 $\boldsymbol{A}$ 中一切主子式都是非负的数．

再证充分性．假定 $\boldsymbol{A}$ 降秩 $(r=r_A<n)$，且其一切主子式都为非负的数．今取 $\boldsymbol{A}$ 之左上角 m 级主子矩阵

$$\boldsymbol{A}_m=\begin{pmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mm} \end{pmatrix},\quad (m=1,\ 2,\ \cdots,\ n)$$

（注意：这里的 $\boldsymbol{A}_m$ 与上面 $\boldsymbol{A}^{(m)}$ 之区别）．于是，令 $\boldsymbol{E}_m$ 表示 m 级单位矩阵，则 $|\lambda\boldsymbol{E}_m+\boldsymbol{A}_m|=\lambda^m+p_1\cdot\lambda^{m-1}+\cdots+p_{m-1}\lambda+p_m$ 中的 p_i 等于 $\boldsymbol{A}_m$ 里面一切 i 级主子式的和（参阅本书第五章第 4 节定理 4），所以据假设可知当 $\lambda>0$ 时，恒有

$$|\lambda E_m+A_m|>0,\ (m=1,\ \cdots,\ n),$$

据定理 2, 可知 $\lambda E+A$ 是正定的(当 $\lambda>0$).

引用这结果, 可用反证法来证明我们需要的结论. 即假定有某一向量 $X\neq 0$ 使 $\overline{X}'AX=-C$, 而 $C>0$; 于是令 $\tau=C/\overline{X}'X$, 显然有 $\tau>0$ 且

$$\overline{X}'(\tau E+A)X=C+\overline{X}'AX=0,$$

这和上面 $\tau E+A$ 应为正定的结论相矛盾. 因之, 对任 X 恒有 $\overline{X}'AX\geqslant 0$, 这证明了 A 是半正定的. 于是, 定理 4 完全获证.

注意: 由于定理 2 之原因, 读者可能会猜想厄米特(实对称)矩阵 $A=(a_{ij})$ 为半正定的充要条件是 $|A|=0$ 及 $|A_m|\geqslant 0(m=1,\cdots,n-1)$, 这个猜想是不正确的, 这只要以二级矩阵 $\begin{pmatrix}0 & 0\\ 0 & -1\end{pmatrix}$ 为例就容易看出. 但是, 却有下面与定理 2 类似的定理, 即

定理 6 设 $A=(a_{ij})$ 是厄米特(实对称)矩阵,

$$A_1=a_{11},\quad A_2=\begin{pmatrix}a_{11} & a_{12}\\ a_{21} & a_{22}\end{pmatrix},\ \cdots,$$

$$A_{n-1}=\begin{pmatrix}a_{11} & \cdots & a_{1,n-1}\\ \vdots & & \vdots\\ a_{n-1,1} & \cdots & a_{n-1,n-1}\end{pmatrix},\ A_n=\begin{pmatrix}a_{11} & \cdots & a_{1n}\\ \vdots & & \vdots\\ a_{n1} & \cdots & a_{nn}\end{pmatrix}.$$

若 $|A_1|>0,\ |A_2|>0,\ \cdots,\ |A_{n-1}|>0,\ |A_n|=0$, 那末 A 是半正定的.

证明 因为 A 是降秩的, 所以只需证明

$$\varphi(x_1,\ \cdots,\ x_n)=\overline{X}'AX\geqslant 0$$

就行了. 用反证法, 如果有某向量 $X\neq 0$, 能使

$$\overline{X}'AX=-C,\quad C>0,$$

则必 $x_n \neq 0$†. 由是 $\tau = C/\bar{x}_n x_n > 0$, 且 $\overline{\boldsymbol{X}}'\boldsymbol{A}\boldsymbol{X} + \tau\bar{x}_n x_n = 0$, 也就是说有 $\boldsymbol{X} \neq 0$, 使得

$$\overline{\boldsymbol{X}}'\boldsymbol{B}\boldsymbol{X} = \boldsymbol{0},$$

这里的 $$\boldsymbol{B} = \begin{pmatrix} a_{11} & \cdots & a_{1,n-1} & a_{1n} \\ \vdots & & \vdots & \vdots \\ a_{n-1,1} & \cdots & a_{n-1,n-1} & a_{n-1,n} \\ a_{n1} & \cdots & a_{n,n-1} & a_{nn} + \tau \end{pmatrix}.$$

另一方面, 因 $\tau > 0$, 所以

$$|\boldsymbol{B}| = |\boldsymbol{A}_n| + \tau|\boldsymbol{A}_{n-1}| > 0,$$

于是再据假设 $|\boldsymbol{A}_1| > 0$, $|\boldsymbol{A}_2| > 0$, $\cdots$, $|\boldsymbol{A}_{n-1}| > 0$, 由定理 2 又知 $\boldsymbol{B}$ 为正定的, 因之 $\overline{\boldsymbol{X}}'\boldsymbol{B}\boldsymbol{X} = 0$ 必然有 $\boldsymbol{X} = \boldsymbol{0}$, 这显然是一个矛盾. 故定理 6 获证.

用行列式方法判定厄米特(或实二次)齐式的值类为正、负定或半正、负定的问题都已经解决了, 剩下的是齐式为不定的问题. 先证下面的

引理 2 厄米特(或实对称)矩阵 $\boldsymbol{A}$ 为正定的充要条件是: i) $\boldsymbol{A}$ 是满秩; 且 ii) $\boldsymbol{A}$ 之一切主子式 $\geqslant 0$.

证明 $\boldsymbol{A}$ 之正定性说明了 $\boldsymbol{A}$ 是满秩的, 故 i) 成立. 又由定理 3 之推论知, $\boldsymbol{A}$ 之一切主子式 > 0, 故也当然是 $\geqslant 0$, 即 ii) 成立. 这就证明了条件的必要性.

条件之充分性的证明与证定理 2 之方法完全类似. 事实

---

† 若 $x_n = 0$, 则一方面有

$$-c = \overline{X}'\boldsymbol{A}\boldsymbol{X} = \sum_1^{n-1} a_{ij}\bar{x}_i x_j = \varphi(x_1, \cdots, x_{n-1}, 0)$$

$$= (\bar{x}_1, \cdots, \bar{x}_{n-1}) \cdot \boldsymbol{A}_{n-1} \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_{n-1} \end{pmatrix},$$

另方面, 由定理 2 又知 $\varphi(x_1, \cdots, x_{n-1}, 0)$ 是正定的, 故不能等于 $-c$.

上，当 $n=1$ 时，$\varphi(x_1)=a_{11}\bar{x}_1x_1$，条件 i) 之意义这时变为 $a_{11}\doteq 0$，条件 ii) 之意义则为 $a_{11}\geqslant 0$，故有 $a_{11}>0$，于是在 $x_1\neq 0$ 时 $a_{11}\bar{x}_1x_1>0$，这说明 $n=1$ 时 $\boldsymbol{A}$ 是正定的.

再归纳地假定 $n-1$ 级矩阵成立．于是，因

$$\boldsymbol{A}_{n-1}=\begin{pmatrix} a_{11} & \cdots & a_{1,n-1} \\ \vdots & & \vdots \\ a_{n-1,1} & \cdots & a_{n-1,n-1} \end{pmatrix}$$

中一切主子式也是 $\boldsymbol{A}$ 的某些主子式，故据条件 ii) 知它们 $\geqslant 0$，因之：当 $\boldsymbol{A}_{n-1}$ 为满秩时，由归纳法之假定则知 $\boldsymbol{A}_{n-1}$ 是正定的，故利用定理 2 可知这时有 $a_{11}>0$，$|\boldsymbol{A}_2|>0$，$\cdots$，$|\boldsymbol{A}_{n-1}|>0$（$\boldsymbol{A}_k$ 之意义参看定理 6），于是条件 i) 与 ii) 之合并又保证了 $|\boldsymbol{A}|=|\boldsymbol{A}_n|>0$，故再次利用定理 2，证明了 $\boldsymbol{A}$ 是正定的.

再考虑 $\boldsymbol{A}_{n-1}$ 降秩．由条件 ii) 知 $\boldsymbol{A}_{n-1}$ 之一切主子式 $\geqslant 0$，故从定理 4 得知 $\boldsymbol{A}_{n-1}$ 是半正定的，因而有 $n-1$ 级满秩矩阵 $\boldsymbol{Q}_1$ 使

$$\bar{\boldsymbol{Q}}_1'\boldsymbol{A}_{n-1}\boldsymbol{Q}_1=\mathrm{diag}\{\underbrace{1,\cdots,1}_{r\text{ 个}},\underbrace{0,\cdots,0}_{n-1-r\text{ 个}}\},\quad r<n-1,$$

若令 $\boldsymbol{Q}=\begin{pmatrix}\boldsymbol{Q}_1 & \\ & 1\end{pmatrix}$，则 $\boldsymbol{Q}$ 为 n 级满秩矩阵，且

$$\boldsymbol{B}=\bar{\boldsymbol{Q}}'\boldsymbol{A}\boldsymbol{Q}=\begin{pmatrix}\bar{\boldsymbol{Q}}_1' & \\ & 1\end{pmatrix}\begin{pmatrix}\boldsymbol{A}_{n-1} & \boldsymbol{S} \\ \bar{\boldsymbol{S}}' & a_{nn}\end{pmatrix}\begin{pmatrix}\boldsymbol{Q}_1 & \\ & 1\end{pmatrix},$$

$$\boldsymbol{S}=\begin{pmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{n-1,n}\end{pmatrix},$$

由简易的计算得

$$B=\begin{pmatrix} \begin{matrix}1 & & \\ & \ddots & \\ & & 1\end{matrix}\ (r\text{个}) & & T \\ & \begin{matrix}0 & & \\ & \ddots & \\ & & 0\end{matrix}\ (n-r-1\text{个}) & \\ \bar{T}' & & a_{nn} \end{pmatrix},$$

$T=\bar{Q}_1'S$ 为 $n-1$ 元的列向量. 故如令

$$T=\begin{pmatrix} b_{1n} \\ b_{2n} \\ \vdots \\ b_{n-1,n} \end{pmatrix},$$

则
$$B=\begin{pmatrix} \begin{matrix}1 & & \\ & \ddots & \\ & & 1\end{matrix}\ (r\text{个}) & & b_{1n} \\ & \begin{matrix}0 & & \\ & \ddots & \\ & & 0\end{matrix}\ (n-r-1\text{个}) & \vdots \\ & & b_{n-1,n} \\ \bar{b}_{1n} \quad \cdots\cdots & \bar{b}_{n-1,n} & a_{nn} \end{pmatrix},$$

令 $X=QY$, 则

$$\begin{aligned}
\varphi(x_1,\ \cdots,\ x_n) &= \bar{X}'AX=\bar{Y}'\bar{Q}'AQY=\bar{Y}'BY \\
&= \bar{y}_1y_1+\cdots+\bar{y}_ry_r+(b_{1n}\bar{y}_1+\cdots+b_{n-1,n}\bar{y}_{n-1})y_n \\
&\quad +(\bar{b}_{1n}y_1+\cdots+\bar{b}_{n-1,n}y_{n-1})\bar{y}_n+a_{nn}\bar{y}_ny_n \\
&= (\bar{y}_1+\bar{b}_{1n}\bar{y}_n)(y_1+b_{1n}y_n)+\cdots \\
&\quad +(\bar{y}_r+\bar{b}_{rn}\bar{y}_n)(y_r+b_{rn}y_n) \\
&\quad +(a_{nn}-\bar{b}_{1n}b_{1n}-\cdots-\bar{b}_{rn}b_{rn})\bar{y}_ny_n \\
&\quad +(\bar{b}_{r+1,n}y_{r+1}+\cdots+\bar{b}_{n-1,n}y_{n-1})\bar{y}_n \\
&\quad +(b_{r+1,n}\bar{y}_{r+1}+\cdots+b_{n-1,n}\bar{y}_{n-1})y_n,
\end{aligned}$$

故满秩变换

$$\begin{cases} z_1 = y_1 \qquad\qquad + b_{1n}y_n, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ z_r = \qquad y_r \qquad + b_{rn}y_n, \\ z_t = \quad y_t (t = r+1, \cdots, n), \end{cases}$$

即
$$\boldsymbol{Y} = \boldsymbol{PZ},\ \boldsymbol{Z} = \begin{pmatrix} z_1 \\ z_2 \\ \vdots \\ z_n \end{pmatrix},$$

$$\boldsymbol{P} = \begin{pmatrix} 1 & \cdots\cdots & -b_{1n} \\ & \ddots & \vdots \\ & 1 \ \cdots\cdots & -b_{rn} \\ & & 1 \ \cdots\cdots \ \vdots \\ & & \ddots \ \vdots \\ & & 1 \end{pmatrix},$$

使
$$\begin{aligned} \varphi(x_1, \cdots, x_n) &= \overline{\boldsymbol{X}}'\boldsymbol{AX} = \overline{\boldsymbol{Y}}'\boldsymbol{BY} = \overline{\boldsymbol{Z}}'\overline{\boldsymbol{P}}'\boldsymbol{BPZ} \\ &= \bar{z}_1 z_1 + \cdots + \bar{z}_r z_r + c z_n z_n + (b_{r+1,n}\bar{z}_{r+1} + \cdots \\ &\quad + b_{n-1,n}\bar{z}_{n-1}) z_n + (\bar{b}_{r+1,n} z_{r+1} + \cdots + \bar{b}_{n-1,n} z_{n-1}) \bar{z}_n, \end{aligned}$$

而 $c = a_{nn} - b_{1n}\bar{b}_{1n} - \cdots - b_{rn}\bar{b}_{rn}$ 为实数. 因之,

$$\overline{\boldsymbol{P}}'\boldsymbol{BP} = \begin{pmatrix} 1 & & & & \\ & \ddots & & & \\ & & 1 & & \\ & & & 0 \cdots\cdots 0 & b_{r+1,n} \\ & & & \vdots \qquad \vdots & \vdots \\ & & & 0 \cdots\cdots 0 & b_{n-1,n} \\ & & & \bar{b}_{r+1,n} \cdots \bar{b}_{n-1,n} & c \end{pmatrix},$$

（其中左上角为 r 阶单位块，标有 r↑）

于是,对行列式说来,就有

$$|\overline{\boldsymbol{P}}'\boldsymbol{BP}| = \begin{vmatrix} 0 \cdots\cdots 0 & b_{r+1,n} \\ \vdots \qquad \vdots & \vdots \\ 0 \cdots\cdots 0 & b_{n-1,n} \\ \bar{b}_{r+1,n} \cdots \bar{b}_{n-1,n} & c \end{vmatrix} \leqslant 0,$$

注意：$r=n-2$ 时，就有 $|\bar{P}'BP| = \begin{vmatrix} 0 & b_{n-1,n} \\ \bar{b}_{n-1,n} & c \end{vmatrix} = -b_{n-1,n}\cdot\bar{b}_{n-1,n}<0$.

但另一方面，$\bar{P}'BP=\bar{P}'\bar{Q}'AQP$，而由于 P，Q 之满秩性及 $|A|>0$，又知 $|\bar{P}'BP|=|\bar{P}|\cdot|P|\cdot|\bar{Q}|\cdot|Q|\cdot|A|>0$，这显然与上面的结果相矛盾．这说明了 A_{n-1} 不可能为降秩的．

于是引理 2 完全获证．

推论 厄米特（或实对称）矩阵 A 为负定的充要条件是：i）A 为满秩，且 ii）A 之偶级主子式 $\geqslant 0$，而一切奇级主子式 $\leqslant 0$．

证明 A 为负定之充要条件是 $-A$ 为正定的，于是，据引理 2 即证得了本推论．

有了引理 2，就容易用行列式的判别法来解决不定的问题，即有

定理 7 厄米特（实对称）矩阵 A 为不定之充要条件是下列二种情况至少有一个成立：

i）A 有一个偶级主子式为负数，

ii）A 有两个奇级主子式，其中一为正数，一为负数．

证明 当条件 i）或 ii）中有一成立时，则由定理 3 之推论，可知 A 非正定亦非负定，又由定理 4 知它亦非半正定，最后据定理 5 又知它也不是半负定的．因之，A 只能是不定的．这就证明了条件的充分性．

反之，设 A 是不定的．若条件 i）不真，即 A 中一切偶级主子式都 $\geqslant 0$，那末分 A 为满秩与降秩两种情况分别讨论于下．

1．A 为满秩．这时，据引理 2，可知 A 中必有一个奇级

主子式为负数，而据引理 2 的推论，知 $\boldsymbol{A}$ 中必有一个奇级主子式为正数．即条件 ii) 必为真．

2．$\boldsymbol{A}$ 是降秩．这时，由定理 4，知 $\boldsymbol{A}$ 至少有一个奇级主子式<0；又由定理 5，知 $\boldsymbol{A}$ 也至少有一个奇级主子式>0，即条件 ii) 成立．

总之，当 $\boldsymbol{A}$ 不定时，不论 $\boldsymbol{A}$ 满秩或降秩，如条件 i) 不真，则条件 ii) 必真．

定理 7 于是完全获证．

注意：定理 7 中两个条件 i) 与 ii) 不是互斥的，即是说可能两个条件都成立；但这两个条件又是独立的，就是说某一个条件成立时，另一个条件也可能不成立．为了说明这点，列举下面的三个例子：

$$\begin{pmatrix}1 & 0 & 1\\ 0 & -1 & 1\\ 1 & 1 & 1\end{pmatrix},\quad \begin{pmatrix}2 & 4 & 2\\ 4 & 2 & 2\\ 2 & 2 & 1\end{pmatrix},\quad \begin{pmatrix}2 & 1 & 1\\ 1 & 2 & -1\\ 1 & -1 & 1\end{pmatrix}.$$

对于第一个例子，定理 7 中 i)、ii) 两条件都成立，但第二个例子只满足条件 i) 却不满足条件 ii)，而第三个例子恰得其反，只适合条件 ii) 而不满足条件 i)．

二次齐式的有定及不定的理论，可应用于数学分析中研究多变量之实函数的极大值与极小值问题．

设 n 元函数 $f(x_1, x_2, \cdots, x_n)$ 在点 $P_0(x_1^0, x_2^0, \cdots, x_n^0)$ 及其近旁连续，并有连续的偏导数．多元函数的泰勒展开式

$$\begin{aligned}&f(x_1^0+\Delta x_1, x_2^0+\Delta x_2, \cdots, x_n^0+\Delta x_n)\\&=f(x_1^0, x_2^0, \cdots, x_n^0)\\&+\left(\Delta x_1\cdot\frac{\partial}{\partial x_1}+\cdots+\Delta x_n\cdot\frac{\partial}{\partial x_n}\right)\cdot f(x_1^0, \cdots, x_n^0)\\&+\frac{1}{2!}\left(\Delta x_1\frac{\partial}{\partial x_1}+\cdots+\Delta x_n\frac{\partial}{\partial x_n}\right)^2\cdot f(x_1^0, \cdots, x_n^0)+R,\end{aligned}$$

可以简写为矩阵之表示式

$$f(P_0+\Delta P_0)=f(P_0)+X'\varphi+\frac{1}{2!}X'AX+R,$$

式中

$$f(P_0)=f(x_1^0,\ \cdots,\ x_n^0),$$

$$f(P_0+\Delta P_0)=f(x_1^0+\Delta x_1,\ \cdots,\ x_n^0+\Delta x_n),$$

$$X=\begin{pmatrix}\Delta x_1\\ \Delta x_2\\ \vdots\\ \Delta x_n\end{pmatrix},\quad \varphi=\begin{pmatrix}\dfrac{\partial f}{\partial x_1}\\ \vdots\\ \dfrac{\partial f}{\partial x_n}\end{pmatrix}\Bigg|_{P_0},$$

$$A=\begin{pmatrix}\dfrac{\partial^2 f}{\partial x_1^2} & \dfrac{\partial^2 f}{\partial x_1\partial x_2} & \cdots & \dfrac{\partial^2 f}{\partial x_1\partial x_n}\\ \cdots\cdots & \cdots\cdots & \cdots\cdots & \cdots\cdots\\ \dfrac{\partial^2 f}{\partial x_n\partial x_1} & \dfrac{\partial^2 f}{\partial x_n\partial x_2} & \cdots & \dfrac{\partial^2 f}{\partial x_n^2}\end{pmatrix}\Bigg|_{P_0},$$

故 A 为实对称矩阵. 在分析学中知道：函数 $f(x_1,\ \cdots,\ x_n)$ 在 P_0 之值为极大或极小，其必要条件是列向量 φ 为零向量，即

$$\left.\frac{\partial f}{\partial x_i}\right|_{P_0}=0\quad (i=1,\ 2,\ \cdots,\ n).$$

若此条件满足，则点 P_0 为临界点，且这时有

$$f(P_0+\Delta P_0)-f(P_0)=\frac{1}{2}X'AX+R,$$

于是，当 Δx_i 足够小时 $(i=1,\ 2,\ \cdots,\ n)$，上式右端之正负号完全由二次齐式 $X'AX$ 来决定. 故若这二次齐式之秩为 n 时，则函数 f 在点 P_0 之近傍的性质可以描述于下：

i) 若 $X'AX$ 为正定的，则 P_0 为一个极小点；

ii) 若 $X'AX$ 为负定的，则 P_0 为一个极大点；

iii) 若 $X'AX$ 为不定的，则 P_0 非极大点也非极小点.

若齐式之秩小于 n 时，为要决定函数 f 在点 P_0 的性态，

则有必要研究余项 R.

现以 $n=2$ 之特例说明之. 设

$$\left.\frac{\partial f}{\partial x_1}\right|_{P_0}=f'_{x_1}(x_1^0,\ x_2^0)=0=f'_{x_2}(x_1^0,\ x_2^0)=\left.\frac{\partial f}{\partial x_2}\right|_{P_0},$$

并设

$$a=\left.\frac{\partial^2 f}{\partial x_1^2}\right|_{P_0},\quad b=\left.\frac{\partial^2 f}{\partial x_1\partial x_2}\right|_{P_0},\quad c=\left.\frac{\partial^2 f}{\partial x_2^2}\right|_{P_0}.$$

于是 $A=\begin{pmatrix} a & b\\ b & c\end{pmatrix}$ 为正定之充要条件是 $a>0$, $ac-b^2>0$, 这时 $f(x_1^0,\ x_2^0)$ 为极小值; A 为负定之充要条件是 $a<0, ac-b^2>0$, 这时 $f(x_1^0,\ x_2^0)$ 为极大值; A 为不定(设 $r_A=2$) 之充要条件是 $ac-b^2<0$, 这时 $f(x_1^0,\ x_2^0)$ 非极大也非极小. 当 $ac-b^2=0$ 时, $r_A<2$, 因之 $f(x,\ y)$ 在点 $P_0(x_1^0,\ x_2^0)$ 处之性态为可疑情形, 即不研究余项 R 就不能决定. 因之可列表于下.

$ac-b^2$	+		−	0
a	+	−	非极值	可疑情形
	极小值	极大值		

上表就是数学分析中大家所熟知的结果.

练习二十九

1. 设 A 为实对称矩阵, $\varphi(x_1,\ \cdots,\ x_n)=\sum a_{ij}x_ix_j=X'AX$ 是 A 所对应的实二次齐式, $\psi(x_1,\ \cdots,\ x_n)=\sum a_{ij}\bar{x}_ix_j=\bar{X}'AX$ 是 A 所对应的厄米特齐式. 证明 φ 和 ψ 属于相同的值类.

2. 证明: 厄米特(实对称)矩阵 $A=(a_{ij})$ 为正定的充要条件是

$$a_{nn}>0,\ \begin{vmatrix} a_{n-1,n-1} & a_{n-1,n}\\ a_{n,n-1} & a_{nn}\end{vmatrix}>0,\ \cdots,\ \begin{vmatrix} a_{11} & \cdots & a_{1n}\\ \vdots & & \vdots\\ a_{n1} & \cdots & a_{nn}\end{vmatrix}>0.$$

3. 设厄米特(实对称)矩阵 $\boldsymbol{A}$ 是降秩的，并且

$$\begin{vmatrix} a_{mm} & \cdots & a_{mn} \\ \vdots & & \vdots \\ a_{nm} & \cdots & a_{nn} \end{vmatrix} > 0, \ (m=2, \cdots, n)$$

试证 $\boldsymbol{A}$ 是半正定的。

4. 设厄米特(实对称)矩阵 $\boldsymbol{A}=(a_{ij})$ 是正定的，$b_1, b_2, \cdots, b_n$ 是任意 n 个不等于零之实数，证明矩阵 $\boldsymbol{B}=(a_{ij}b_ib_j)$ 也是正定的。

5. 设有两个厄米特齐式 $\overline{\boldsymbol{X}}'\boldsymbol{A}\boldsymbol{X}$ 与 $\overline{\boldsymbol{X}}'\boldsymbol{B}\boldsymbol{X}$，后者是满秩的，并且它们能够同时通过一个满秩变换变成典型式 $a_1\bar{x}_1x_1+\cdots+a_n\bar{x}_nx_n$ 与 $b_1\bar{x}_1x_1+\cdots+b_n\bar{x}_nx_n$，试证 $|\boldsymbol{A}-\lambda\boldsymbol{B}|=0$ 的 n 个根为实数，并求出它们。由是再证明 x^2+2xy 与 $2xy+y^2$ 不能通过同一个满秩变换变成平方和。

6. 若 $\boldsymbol{A}$ 是正定的，试证其逆 $\boldsymbol{A}^{-1}$ 与伴随 $\boldsymbol{A}^*$ 也都是正定的。

7. 设 $\boldsymbol{A}=\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}$ 是正定的，并令 $\boldsymbol{A}_{n-1}=\begin{pmatrix} a_{11} & \cdots & a_{1,n-1} \\ \vdots & & \vdots \\ a_{n-1,1} & \cdots & a_{n-1,n-1} \end{pmatrix}$，

证明：
$$a_{nn}\cdot|\boldsymbol{A}_{n-1}|>\overline{\boldsymbol{X}}'\boldsymbol{A}_{n-1}^*\boldsymbol{X},$$

但
$$\boldsymbol{X}=\begin{pmatrix} x_1 \\ \vdots \\ x_{n-1} \end{pmatrix}=\begin{pmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{n-1,n} \end{pmatrix}.$$

8. 设 $\varphi(x_1, \cdots, x_n)=\sum a_{ij}x_ix_j$ 是不定的实二次齐式，它在点 $P(y_1, \cdots, y_n)$ 的值为正数：$\varphi(y_1, \cdots, y_n)>0$；而在点 $Q(z_1, \cdots, z_n)$ 的值是负数：$\varphi(z_1, \cdots, z_n)<0$。证明：有两个且只有两个和 P、Q 线性相关的点 $R(u_1, \cdots, u_n)$、$S(v_1, \cdots, v_n)$，使 $\varphi(x_1, \cdots, x_n)$ 之值等于零，并证 R 与 S 是线性无关的。

9. 证明实二次齐式 $n(x_1^2+\cdots+x_n^2)-(x_1+\cdots+x_n)^2$ 是半正定的，并证当且仅当 $x_1=\cdots=x_n$ 时其值才为零。

10. 当 λ 为何值时，四变量的实二次齐式

$$\lambda(x^2+y^2+z^2)+2xy-2yz+2zx+t^2$$

为正定的？并讨论 $\lambda=2$ 之情况。

11. 设 $\boldsymbol{A}=(a_{ij})$ 是一个 n 级的正定厄米特矩阵。试证：$|\boldsymbol{A}|\leqslant a_{11}a_{22}\cdots a_{nn}$，且等号"="成立的充要条件是 $\boldsymbol{A}$ 为对角矩阵。

12. 设 $\boldsymbol{A}$ 是正定的，那末只要 $\boldsymbol{Q}$ 是满秩，$\overline{\boldsymbol{Q}}'\boldsymbol{A}\boldsymbol{Q}$ 也是正定的。同样，对于 $\boldsymbol{A}$ 为负定，半正(负)定或不定时，$\overline{\boldsymbol{Q}}'\boldsymbol{A}\boldsymbol{Q}$ 也有类似的结果。

13. 设 $\varphi(x_1, \cdots, x_n)=\sum\limits_{i,j=1}^{n} a_{ij}x_i\bar{x}_j$ 为正定厄米特齐式(或实二次齐式)，则有如下的满秩变换(或实满秩变换)：

$$y_i=x_i+a_{i,i+1}x_{i+1}+\cdots+a_{in}x_n,\ (i=1,\ 2,\ \cdots,\ n)$$

使 φ 变形为 $\sum\limits_{i=1}^{n}c_iy_i\bar{y}_i$（或 $\sum\limits_{i=1}^{n}c_iy_i^2$），$c_i>0$, $i=1,\ 2,\ \cdots,\ n$.

14. 设复系数方程 $x^n+a_1x^{n-1}+\cdots+a_{n-1}x+a_n=0$ 之 n 个根为 λ_1、λ_2、$\cdots$、λ_n，并令 $s_k=\lambda_1^k+\lambda_2^k+\cdots+\lambda_n^k$. 试证这方程之互异根的个数，等于下述矩阵的秩：

$$\boldsymbol{S}=\begin{pmatrix} S_0 & S_1 & \cdots & S_{n-1}\\ S_1 & S_2 & \cdots & S_n\\ \vdots & \vdots & & \vdots\\ S_{n-1} & S_n & \cdots & S_{2n-2}\end{pmatrix}.$$

15. 实系数 n 次方程的 n 个根互异的充要条件是 $\boldsymbol{S}$ 为正定对称矩阵（$\boldsymbol{S}$ 的意义参见上题）.

16. 实对称矩阵 $\boldsymbol{A}$ 的所有特征值在区间 $[a,\ b]$ 内的充要条件是对任 $\lambda_0<a$, $\boldsymbol{A}-\lambda_0\boldsymbol{E}$ 是正定的；而对任 $\lambda_0>b$, $\boldsymbol{A}-\lambda_0\boldsymbol{E}$ 为负定的.

17. 设 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 都是实对称矩阵，证明：若 $\boldsymbol{A}$ 的特征值在区间 $[a,\ b]$ 内，$\boldsymbol{B}$ 的特征值在区间 $[c,\ d]$ 内，则 $\boldsymbol{A}+\boldsymbol{B}$ 的特征值必在区间 $[a+c,\ b+d]$ 内.

第6节　厄米特齐式的分解因式

对比第 4 节与第 1、2 节，可以发现，实二次齐式的主要结果或性质对厄米特齐式全都成立，而且所用的方法也完全一样．换言之，关于实二次齐式的所用方法全都可以照样推广到厄米特齐式上去．可是，在第 2 节里曾介绍了实二次齐式的因式分解，而对厄米特齐式的因式分解怎样呢？本节来讨论这个问题.

首先，厄米特齐式如可分解为一次因式的积，则两个一次因式之形状是什么？当然，如能分解，则每个一次因式的普遍形状不外乎是

$$a_1x_1+a_2x_2+\cdots+a_nx_n+\alpha_1\bar{x}_1+\alpha_2\bar{x}_2+\cdots+\alpha_n\bar{x}_n.$$

因之，若厄米特齐式 $\sum a_{ij}\bar{x}_ix_j$ 可分解，就应为

$$\sum a_{ij}\bar{x}_i x_j=(a_1x_1+\cdots+a_nx_n+\alpha_1\bar{x}_1+\cdots+\alpha_n\bar{x}_n)$$
$$\cdot(b_1x_1+\cdots+b_nx_n+\beta_1\bar{x}_1+\cdots+\beta_n\bar{x}_n)$$

形；故当 $a_1, \cdots, a_n$ 中有一不为零（如 $a_1\neq0$）时，则因展开右端后得知 x_1^2 之系数为 a_1b_1，而左端又不含 x_1^2，所以 $a_1b_1=0$，既 $a_1\neq0$ 故必 $b_1=0$；这样，当 $i=2, \cdots, n$ 时，右端 x_1x_i 之系数 $=a_1b_i+a_ib_1=a_1b_i$，左端不含 x_1x_i，故又有 $a_1b_i=0$，从而 $b_i=0$ $(i=2, \cdots, n)$．于是，

$$\sum a_{ij}\bar{x}_i x_j=(a_1x_1+\cdots+a_nx_n+\alpha_1\bar{x}_1+\cdots+\alpha_n\bar{x}_n)(\beta_1\bar{x}_1+\cdots+\beta_n\bar{x}_n).$$

然 $\sum a_{ij}\bar{x}_i x_j$ 为非零的齐式，故 $\beta_1, \cdots, \beta_n$ 中必有一个 $\neq0$，如令 $\beta_1\neq0$，则展开上式右端得 $\bar{x}_1^2$ 之系数 $=\alpha_1\beta_1$，故同上述的理由可知 $\alpha_1=0$，随而再比较 $\bar{x}_1\bar{x}_i(i\neq1)$ 之系数又得 $0=\alpha_1\beta_i+\alpha_i\beta_1=\alpha_i\beta_1$，故 $\alpha_i=0(i\neq1)$．因之得到了

$$\sum a_{ij}\bar{x}_i x_j=(a_1x_1+\cdots+a_nx_n)(\beta_1\bar{x}_1+\cdots+\beta_n\bar{x}_n).$$

这说明了厄米特齐式若能分解，所分解之因式必为上形．为对称计，将分解之形改写为

$$\sum a_{ij}\bar{x}_i x_j=(\bar{\alpha}_1\bar{x}_1+\cdots+\bar{\alpha}_n\bar{x}_n)(\beta_1x_1+\cdots+\beta_nx_n). \qquad (15)$$

再与第 2 节中一样，分两种情况讨论于下．

i) $\alpha_1x_1+\cdots+\alpha_nx_n$ 与 $\beta_1x_1+\cdots+\beta_nx_n$ 线性相关．

这时，有 ρ 使 $\beta_i=\rho\alpha_i(i=1, \cdots, n)$，于是就从(15)式得到

$$\sum a_{ij}\bar{x}_i x_j=\rho\cdot\left(\sum_{s=1}^n\bar{\alpha}_s\bar{x}_s\right)\cdot\left(\sum_{s=1}^n\alpha_sx_s\right),$$

因之满秩变换 $y_1=\sum\alpha_sx_s$, $y_j=x_j(j=2, \cdots, n)$（因 $\alpha_1, \cdots, \alpha_n$ 不全为 0，故不失一般性可令 $\alpha_1\neq0$）使

$$\sum a_{ij}\bar{x}_i x_j=\rho\bar{y}_1y_1,$$

故 $r_A=1$，同时也知道 ρ 为实数．

ii) $\alpha_1x_1+\cdots+\alpha_nx_n$ 与 $\beta_1x_1+\cdots+\beta_nx_n$ 线性无关．

这时，不妨令 $\begin{vmatrix} \alpha_1 & \beta_1 \\ \alpha_2 & \beta_2 \end{vmatrix} \neq 0$，于是满秩变换 $y_1 = \sum \alpha_i x_i$，$y_2 = \sum \beta_i x_i$，$y_t = x_t (t=3, \cdots, n)$ 使(15)变为

$$\sum a_{ij}\bar{x}_i x_j = \bar{y}_1 y_2.$$

但因 $\sum a_{ij}\bar{x}_i x_j$ 之共轭等于它自身(因 $\sum a_{ij}\bar{x}_i x_j$ 为实数)，故 $\bar{y}_1 y_2 = y_1 \bar{y}_2$，于是有

$$\sum a_{ij}\bar{x}_i x_j = \frac{1}{2}(\bar{y}_1 y_2 + y_1 \bar{y}_2).$$

再作满秩变换

$$y_1 = z_1 - z_2,\ y_2 = z_1 + z_2,\ y_t = z_t (t=3, \cdots, n),$$

即得
$$\sum a_{ij}\bar{x}_i x_j = \bar{z}_1 z_1 - \bar{z}_2 z_2,$$

证明了 $r_A = 2$，正负号差 $s = 0$.

反之，如果 $r_A = 1$，则典型式应为

$$\sum a_{ij}\bar{x}_i x_j = \lambda \bar{y}_1 y_1$$

之形状，故若 $y_1 = \alpha_1 x_1 + \cdots + \alpha_n x_n$，则

$$\sum a_{ij}\bar{x}_i x_j = \lambda(\bar{\alpha}_1 \bar{x}_1 + \cdots + \bar{\alpha}_n \bar{x}_n)(\alpha_1 x_1 + \cdots + \alpha_n x_n).$$

但若 $r_A = 2$，$s = 0$，则这时典型式为

$$\sum a_{ij}\bar{x}_i x_j = \bar{y}_1 y_1 - \bar{y}_2 y_2$$

形；如令
$$y_1 = \lambda_1 x_1 + \cdots + \lambda_n x_n,$$
$$y_2 = \mu_1 x_1 + \cdots + \mu_n x_n,$$

则因 y_1，y_2 线性无关，故与第2节中的证法完全一样，知

$$z_1 = y_1 + y_2 = (\lambda_1 + \mu_1)x_1 + \cdots + (\lambda_n + \mu_n)x_n,$$
$$z_2 = y_1 - y_2 = (\lambda_1 - \mu_1)x_1 + \cdots + (\lambda_n - \mu_n)x_n$$

也线性无关．于是，因 $\frac{1}{2}(\bar{z}_1 z_2 + z_1 \bar{z}_2) = \bar{y}_1 y_1 - \bar{y}_2 y_2$，故

$$\sum a_{ij}\bar{x}_i x_j = \frac{1}{2}(\bar{z}_1 z_2 + z_1 \bar{z}_2).$$

可是，到了这一步就再不敢保证有 $\sum a_{ij}\bar{x}_i x_j = \bar{z}_1 z_2$，因为不见得

有 $\bar{z}_1 z_2 = \bar{z}_2 z_1$ 的缘故，所以 $\sum a_{ij}\bar{x}_i x_j$ 到底能否分解，值得考虑．而且，实际上两个变量 x_1，x_2 的厄米特齐式 $\bar{x}_1 x_2 + \bar{x}_2 x_1$（秩为 2，正负号差 $=0$）显然是不可分解的†．

于是，若完全按照第 2 节的方法平行地去推，那末我们只能说：如厄米特齐式能分解，则 $r_A=1$，或 $r_A=2, s=0$；反之，$r_A=1$ 时又确可保证能分解，而 $r_A=2$ 及 $s=0$ 时则成了疑难问题．仔细检查，发现第 ii) 种情况不会发生（当厄米特齐式 $\sum a_{ij}\bar{x}_i x_j$ 能分解如(15)之形状时）．这是因为，从 (15) 之展开得 $\bar{\alpha}_1\beta_1$ 与 $\bar{\alpha}_2\beta_2$ 应分别等于 a_{11} 与 a_{22}，而为实数，故有 $\beta_1 = a\alpha_1$，$\beta_2 = b\alpha_2$（a，b 均为实数）；又因 $\bar{\alpha}_1\beta_2 = a_{12}$，$\bar{\alpha}_2\beta_1 = a_{21}$ 互为共轭的，故代入之即知 $\bar{\alpha}_1\beta_2 = b\bar{\alpha}_1\alpha_2$ 与 $\bar{\alpha}_2\beta_1 = a\alpha_1\bar{\alpha}_2$ 共轭，因之 $b\bar{\alpha}_1\alpha_2 = \overline{(a\alpha_1\bar{\alpha}_2)} = \bar{a}\bar{\alpha}_1\alpha_2 = a\bar{\alpha}_1\alpha_2$，即 $(b-a)\cdot\bar{\alpha}_1\alpha_2 = 0$，故或 $(b-a)\alpha_2 = 0$，或 $\bar{\alpha}_1 = 0$ 随而 $\alpha_1 = 0$，于是不论 $(b-a)\alpha_2 = 0$ 或 $\alpha_1 = 0$，均有

$$\begin{vmatrix} \alpha_1 & \alpha_2 \\ \beta_1 & \beta_2 \end{vmatrix} = \begin{vmatrix} \alpha_1 & \alpha_2 \\ a\alpha_1 & b\alpha_2 \end{vmatrix} = (b-a)\alpha_2\cdot\alpha_1 = 0;$$

同理，$\begin{vmatrix} \alpha_k & \alpha_l \\ \beta_k & \beta_l \end{vmatrix} = 0\ (k \neq l)$．这说明 $\alpha_1 x_1 + \cdots + \alpha_n x_n$ 与 $\beta_1 x_1 + \cdots + \beta_n x_n$ 必线性相关，这就说明了第 ii) 种情况不会发生．这是厄米特齐式与实二次齐式在分解因式时之不同之处．故厄米特齐式可分解为一次因式之积的充要条件是 $r_A = 1$．这结论也可以用另外的方法来证明于下：

若 $\sum a_{ij}\bar{x}_i x_j$ 可分解为(15)之形，则易看出 $a_{ij} = \bar{\alpha}_i\beta_j$，故得知

---

† 因若能分解，则必 $\bar{x}_1 x_2 + \bar{x}_2 x_1 = (\bar{\alpha}_1\bar{x}_1 + \bar{\alpha}_2\bar{x}_2)(\beta_1 x_1 + \beta_2 x_2) = \bar{\alpha}_1\beta_1\bar{x}_1 x_1 + \bar{\alpha}_2\beta_2\bar{x}_2 x_2 + \bar{\alpha}_1\beta_2\bar{x}_1 x_2 + \bar{\alpha}_2\beta_1\bar{x}_2 x_1$，故有 $\bar{\alpha}_1\beta_1 = 0 = \bar{\alpha}_2\beta_2$，$\bar{\alpha}_1\beta_2 = 1 = \bar{\alpha}_2\beta_1$，因之一方面有 $\bar{\alpha}_1\bar{\alpha}_2\beta_1\beta_2 = 0$，而同时又有 $\bar{\alpha}_1\bar{\alpha}_2\beta_1\beta_2 = 1$，这显非所许．

$$\begin{vmatrix} a_{ij} & a_{ik} \\ a_{lj} & a_{lk} \end{vmatrix} = \begin{vmatrix} \bar{\alpha}_i\beta_j & \bar{\alpha}_i\beta_k \\ \bar{\alpha}_l\beta_j & \bar{\alpha}_l\beta_k \end{vmatrix} = \bar{\alpha}_i\bar{\alpha}_l \begin{vmatrix} \beta_j & \beta_k \\ \beta_j & \beta_k \end{vmatrix} = 0,$$

即 $r_A = 1$.

附注：从这个直接的证明方法，即知前述的第 ii) 种情况不会发生，这样还简单得多. 但应注意这方法不能用在实二次齐式 $\sum a_{ij}x_ix_j$ 上，因为从 $\sum a_{ij}x_ix_j = (\alpha_1x_1 + \cdots + \alpha_nx_n)\cdot(\beta_1x_1 + \cdots + \beta_nx_n)$，只能说 $2a_{ij} = a_{ij} + a_{ji} = \alpha_i\beta_j + \alpha_j\beta_i$，而 $\alpha_i\beta_j$ 与 $\alpha_j\beta_i$ 不见得相等，因之有可能 $a_{ij} \neq \alpha_i\beta_j$.

反之，当 $r_A = 1$ 时，则按前述的方法知 $\sum a_{ij}\bar{x}_ix_j$ 的典型式为 $\lambda\bar{y}_1y_1$，由之得分解之形为

$$\sum a_{ij}\bar{x}_ix_j = \lambda(\bar{\alpha}_1\bar{x}_1 + \cdots + \bar{\alpha}_n\bar{x}_n)(\alpha_1x_1 + \cdots + \alpha_nx_n).$$

或直接证明于下(即不用典型式)：

因这时有 $2n$ 个数 $c_1, \cdots, c_n, d_1, \cdots, d_n$ 使得 $a_{ij} = c_id_j$ (参看本书练习六的第 6 题)，因之

$$\sum a_{ij}\bar{x}_ix_j = (c_1\bar{x}_1 + \cdots + c_n\bar{x}_n)(d_1x_1 + \cdots + d_nx_n).$$

但用这样直接的方法不能立即断定两个一次因式的形状，这是要留心的.

总括上述，证得了下面的

定理 厄米特齐式 $\sum a_{ij}\bar{x}_ix_j$ 能分解为两个一次因式之积的充要条件是厄米特矩阵 $A = (a_{ij})$ 的秩 r_A 等于 $1 (r_A = 1)$；这时两个因式的形状是

$$\bar{\alpha}_1\bar{x}_1 + \cdots + \bar{\alpha}_n\bar{x}_n \text{ 与 } \alpha_1x_1 + \cdots + \alpha_nx_n.$$

第七章　欧氏空间

前已指出(参见本书第六章)：若把二次齐式化为典型式的这种演化，看作是解析几何中把二次有心曲线（或曲面）化为典型式的推广，那就需要引进与前章开头所说的变换(2)类似的所谓直交变换．用直变变换把二次齐式化为典型式的理论是线性代数中最重要的一个问题，叫做主轴问题．由于直交变换这个概念牵涉到空间的度量性质，因此我们先在仿射空间里引进度量概念而使之成为欧氏空间．本章研究欧氏空间，以主轴问题为中心问题．并且，本章只讨论实空间，复空间留在下章去讨论．

第1节　欧氏空间的概念

在解析几何里，向量的长和两个向量的夹角，是度量的两个基本概念，但这两者都和向量的数积紧密相关，即向量的长等于这向量和它自己的数积的平方根，二向量间夹角的余弦等于它们的数积除以它们的长之乘积．用数学式子来表示就是

$$\|\boldsymbol{x}\| = \sqrt{(\boldsymbol{x},\ \boldsymbol{x})}, \qquad \cos\theta = \frac{(\boldsymbol{x},\ \boldsymbol{y})}{\|\boldsymbol{x}\|\cdot\|\boldsymbol{y}\|},$$

式中符号$(\boldsymbol{x},\ \boldsymbol{y})$表示向量$\boldsymbol{x}$，$\boldsymbol{y}$之数积，$\|\boldsymbol{x}\|$为向量$\boldsymbol{x}$的长，$\theta = (\widehat{\boldsymbol{x},\ \boldsymbol{y}})$表示$\boldsymbol{x}$，$\boldsymbol{y}$间的夹角．众所周知，关于数积，还有下述基本性质：$(\boldsymbol{x},\ \boldsymbol{y}) = (\boldsymbol{y},\ \boldsymbol{x})$；$(\lambda\boldsymbol{x},\ \boldsymbol{y}) = \lambda(\boldsymbol{x},\ \boldsymbol{y})$；$(\boldsymbol{x},\ \boldsymbol{x}) \geqslant$

0；当且仅当 $\boldsymbol{x}=\mathbf{0}$ 时才有 $(\boldsymbol{x}, \boldsymbol{x})=0$；以及 $(\boldsymbol{x}_1+\boldsymbol{x}_2, \boldsymbol{y})=(\boldsymbol{x}_1, \boldsymbol{y})+(\boldsymbol{x}_2, \boldsymbol{y})$ 等．在仿射空间中，也就可以先规定数积的意义，然后再用数积来定义向量的长及二个向量的夹角．本来，在解析几何里，是先有向量的长及向量间的夹角这两个具体的东西，然后根据它们才抽象得到所谓向量之数积这个概念；可是，在 n 维空间 $\mathfrak{R}_n(n>3)$ 内，已没有直觉的几何意义，欲明确向量的长及夹角，须从代数式子去着手，也就是先要规定数积的意义，并把直觉空间中数积的上述这些性质作为公理来要求而使 $\mathfrak{R}_n$ 成为有度量性质的空间，叫做欧氏空间，即：

定义　对于仿射(实)空间 $\mathfrak{R}$ 中任二个向量 $\boldsymbol{x}$、$\boldsymbol{y}$，恒有一个确定的方法使之对应于一个相应的实数，记为 $(\boldsymbol{x}, \boldsymbol{y})$，并叫它为 $\boldsymbol{x}$、$\boldsymbol{y}$ 之数积†，那末，当下列四个条件（或公理）能满足时，就叫 $\mathfrak{R}$ 为欧氏空间．

1° $(\boldsymbol{x}, \boldsymbol{y})=(\boldsymbol{y}, \boldsymbol{x})$；

2° $(\lambda\boldsymbol{x}, \boldsymbol{y})=\lambda(\boldsymbol{x}, \boldsymbol{y})$，$\lambda$ 是任何实数；

3° $(\boldsymbol{x}_1+\boldsymbol{x}_2, \boldsymbol{y})=(\boldsymbol{x}_1, \boldsymbol{y})+(\boldsymbol{x}_2, \boldsymbol{y})$；

4° $(\boldsymbol{x}, \boldsymbol{x})\geqslant 0$，而等于零是在且仅在 $\boldsymbol{x}=\mathbf{0}$ 时．

欧氏空间内向量 $\boldsymbol{x}$ 的长定义为 $\|\boldsymbol{x}\|=\sqrt{(\boldsymbol{x}, \boldsymbol{x})}$，两个非零向量 $\boldsymbol{x}$, $\boldsymbol{y}$ 间的夹角 $\theta=(\widehat{\boldsymbol{x}, \boldsymbol{y}})$ 的余弦定义为 $\cos\theta=\dfrac{(\boldsymbol{x}, \boldsymbol{y})}{\|\boldsymbol{x}\|\cdot\|\boldsymbol{y}\|}$.

从定义易知 $(\boldsymbol{x}, \lambda\boldsymbol{y})=\lambda(\boldsymbol{x}, \boldsymbol{y})$ 与 $(\boldsymbol{x}, \boldsymbol{y}+\boldsymbol{z})=(\boldsymbol{x}, \boldsymbol{y})+(\boldsymbol{x}, \boldsymbol{z})$．还要指出：非零向量 $\boldsymbol{x}$ 的长 $\|\boldsymbol{x}\|$ 总是正数，只有零向量的长才为零．又，零向量 $\mathbf{0}$ 与任何向量 $\boldsymbol{x}$ 的数积等于 0，即 $(\boldsymbol{x}, \mathbf{0})=0$（这是因为 $(\boldsymbol{x}, \mathbf{0})=(\boldsymbol{x}, \boldsymbol{y}-\boldsymbol{y})=(\boldsymbol{x}, \boldsymbol{y})-$

† 也就是说，数积 $(\boldsymbol{x}, \boldsymbol{y})$ 为向量 $\boldsymbol{x}$, $\boldsymbol{y}$ 的函数，而自变量是向量，函数值是实数．

$(\boldsymbol{x}, \boldsymbol{y})=0$). 长为 1 的向量叫做单位向量. 因为 $\|\lambda\boldsymbol{x}\|=\sqrt{(\lambda\boldsymbol{x}, \lambda\boldsymbol{x})}=|\lambda|\cdot\sqrt{(\boldsymbol{x}, \boldsymbol{x})}$, 所以

$$\left\|\frac{\boldsymbol{x}}{\sqrt{(\boldsymbol{x}, \boldsymbol{x})}}\right\|=\frac{1}{\sqrt{(\boldsymbol{x}, \boldsymbol{x})}}\cdot\sqrt{(\boldsymbol{x}, \boldsymbol{x})}=1,$$

这就是说 $\boldsymbol{x}$ 虽不是单位向量, 但 $\dfrac{\boldsymbol{x}}{\sqrt{(\boldsymbol{x}, \boldsymbol{x})}}$ 确是单位向量, 这种由 $\boldsymbol{x}$ 求单位向量 $\dfrac{\boldsymbol{x}}{\sqrt{(\boldsymbol{x}, \boldsymbol{x})}}$ 的过程叫做 $\boldsymbol{x}$ 的标准化.

我们知道 $|\cos\theta|\leqslant 1$, 为了说明上面定义的两个非零的向量 $\boldsymbol{x}, \boldsymbol{y}$ 间的夹角 θ 的余弦的合理性, 就需要证明下面的不等式:

$$(\boldsymbol{x}, \boldsymbol{y})^2\leqslant\|\boldsymbol{x}\|^2\cdot\|\boldsymbol{y}\|^2=(\boldsymbol{x}, \boldsymbol{x})\cdot(\boldsymbol{y}, \boldsymbol{y}). \tag{1}$$

证明 不论 ξ、η 是任何的实数, 由 4° 可知

$$(\xi\boldsymbol{x}+\eta\boldsymbol{y}, \xi\boldsymbol{x}+\eta\boldsymbol{y})\geqslant 0; \tag{2}$$

但由 1°～3°, 可以得到

$$(\xi\boldsymbol{x}+\eta\boldsymbol{y}, \xi\boldsymbol{x}+\eta\boldsymbol{y})$$
$$=\xi^2\cdot(\boldsymbol{x}, \boldsymbol{x})+2\xi\eta\cdot(\boldsymbol{x}, \boldsymbol{y})+\eta^2\cdot(\boldsymbol{y}, \boldsymbol{y}),$$

于是 (2) 式之意义即是说上式右边是实变量 ξ, η 的正定或半正定二次齐式, 因之

$$\begin{vmatrix}(\boldsymbol{x}, \boldsymbol{x}) & (\boldsymbol{x}, \boldsymbol{y})\\ (\boldsymbol{x}, \boldsymbol{y}) & (\boldsymbol{y}, \boldsymbol{y})\end{vmatrix}\geqslant 0,$$

即证得了 (1) 式.

注意: 即令 $\boldsymbol{x}, \boldsymbol{y}$ 中有为零向量的, (1) 式仍然成立, 因这时 (1) 之左右两边都等于零.

另外, 还要说明的是: 今后, 两个非零向量 $\boldsymbol{x}, \boldsymbol{y}$ 间的夹角 θ, 限制在 $0\leqslant\theta\leqslant\pi$ 的范围内.

(1) 式叫做柯西不等式, 由上面的证明知道: (1) 中等号成立的充要条件是适合 $\xi'\boldsymbol{x}+\eta'\boldsymbol{y}=\boldsymbol{0}$ 的二个不为 0 之实数 ξ',

η' 必定存在，也就是说 $\boldsymbol{x}$ 与 $\boldsymbol{y}$ 是线性相关的，即 $\boldsymbol{y}=\lambda\boldsymbol{x}$，这时 $\cos\theta=\dfrac{\lambda\cdot(\boldsymbol{x},\ \boldsymbol{x})}{\|\boldsymbol{x}\|\cdot\sqrt{\lambda^2}\,\|\boldsymbol{x}\|}=\dfrac{\lambda}{|\lambda|}=\pm1$，所以当 $\lambda>0$ 时 $\theta=0$，而当 $\lambda<0$ 时 $\theta=\pi$；但不论是 $\lambda>0$ 或 $\lambda<0$，这时叫 $\boldsymbol{x}$ 与 $\boldsymbol{y}$ 平行，用 $\boldsymbol{x}/\!/\boldsymbol{y}$ 来表示.

又，当 $\theta=\dfrac{\pi}{2}$ 时，叫 $\boldsymbol{x}$ 与 $\boldsymbol{y}$ 垂直(或直交)，而用 $\boldsymbol{x}\perp\boldsymbol{y}$ 来表示.

所以，$\boldsymbol{x}/\!/\boldsymbol{y}$ 的充要条件是 $\boldsymbol{x}=\lambda\boldsymbol{y}$ (或 $\boldsymbol{y}=\lambda\boldsymbol{x}$)，而 $\boldsymbol{x}\perp\boldsymbol{y}$ 的充要条件是 $(\boldsymbol{x},\ \boldsymbol{y})=0$.

［例 1］ 由一切 n 个实数组 $\boldsymbol{x}=(x_1,\ x_2,\ \cdots,\ x_n)$ 所成的空间 R_n 中，定义二向量 $\boldsymbol{x}=(x_1,\ \cdots,\ x_n)$ 和 $\boldsymbol{y}=(y_1,\ \cdots,\ y_n)$ 的数积为 $(\boldsymbol{x},\ \boldsymbol{y})=x_1y_1+x_2y_2+\cdots+x_ny_n$. 容易验证这时公理 $1^\circ\sim4^\circ$ 都满足. 柯西不等式(1)在这时是

$$\left(\sum_{i=1}^{n}x_iy_i\right)^2\leqslant\sum_{i=1}^{n}x_i^2\cdot\sum_{i=1}^{n}y_i^2.$$

此即第四章第 8 节中曾经指出过的一个很重要的不等式.

［例 2］ 在变量 x 的次数不大于 n 的所有多项式构成的 n 维空间 $P[x]_n$ 中，两个向量 $f(x)$，$g(x)$ 的数积如果定义为 $(f(x),\ g(x))=\int_a^b f(x)g(x)dx$，则易验证这时公理 $1^\circ\sim4^\circ$ 都满足. 于是，这时柯西不等式(1)则为

$$\left(\int_a^b f(x)g(x)dx\right)^2\leqslant\int_a^b[f(x)]^2dx\cdot\int_a^b[g(x)]^2dx.$$

有了柯西不等式，容易证明

$$\big|\|\boldsymbol{x}\|-\|\boldsymbol{y}\|\big|\leqslant\|\boldsymbol{x}\pm\boldsymbol{y}\|\leqslant\|\boldsymbol{x}\|+\|\boldsymbol{y}\|; \tag{3}$$

$$\|\boldsymbol{x}\pm\boldsymbol{y}\|^2=\|\boldsymbol{x}\|^2+\|\boldsymbol{y}\|^2,\ \text{当}\ \boldsymbol{x}\perp\boldsymbol{y}\ \text{时}. \tag{4}$$

事实上，由柯西不等式 $(\boldsymbol{x},\ \boldsymbol{y})^2\leqslant\|\boldsymbol{x}\|^2\cdot\|\boldsymbol{y}\|^2$，即得

$$-\|\boldsymbol{x}\|\cdot\|\boldsymbol{y}\|\leqslant(\boldsymbol{x},\ \boldsymbol{y})\leqslant\|\boldsymbol{x}\|\cdot\|\boldsymbol{y}\|.$$

于是从

$$\begin{aligned}\|\boldsymbol{x}\pm\boldsymbol{y}\|^2=(\boldsymbol{x}\pm\boldsymbol{y},\ \boldsymbol{x}\pm\boldsymbol{y})&=(\boldsymbol{x},\ \boldsymbol{x})+(\boldsymbol{y},\ \boldsymbol{y})\\&\pm2\cdot(\boldsymbol{x},\ \boldsymbol{y})=\|\boldsymbol{x}\|^2+\|\boldsymbol{y}\|^2\pm2\cdot(\boldsymbol{x},\ \boldsymbol{y}),\end{aligned}$$

而利用 $\boldsymbol{x}\perp\boldsymbol{y}$（即 $(\boldsymbol{x},\ \boldsymbol{y})=0$）马上就得到了(4)式，一般地，

$$\begin{aligned}\|\boldsymbol{x}\pm\boldsymbol{y}\|^2=\|\boldsymbol{x}\|^2+\|\boldsymbol{y}\|^2\pm2\cdot(\boldsymbol{x},\ \boldsymbol{y})&\\&\begin{cases}\leqslant\|\boldsymbol{x}\|^2+\|\boldsymbol{y}\|^2+2\cdot\|\boldsymbol{x}\|\cdot\|\boldsymbol{y}\|,\\\geqslant\|\boldsymbol{x}\|^2+\|\boldsymbol{y}\|^2-2\cdot\|\boldsymbol{x}\|\cdot\|\boldsymbol{y}\|,\end{cases}\end{aligned}$$

即 $(\|\boldsymbol{x}\|-\|\boldsymbol{y}\|)^2\leqslant\|\boldsymbol{x}\pm\boldsymbol{y}\|^2\leqslant(\|\boldsymbol{x}\|+\|\boldsymbol{y}\|)^2$，

故(3)式成立.

在几何里，(3)之几何意义是三角形的一边小于另二边的和而大于另二边的差，(4)之几何意义就是直角三角形两直角边平方的和等于斜边的平方.

我们还知道：在解析几何里，两点 $\boldsymbol{x}$ 与 $\boldsymbol{y}$ 间的距离定义为向量 $\boldsymbol{x}-\boldsymbol{y}$ 的长，在 n 维欧氏空间里，我们也用公式

$$d=\|\boldsymbol{x}-\boldsymbol{y}\|$$

来定义 $\boldsymbol{x}$ 和 $\boldsymbol{y}$ 间的距离.

我们知道，在研究 n 维仿射空间时，基底是起主要作用的，譬如在 n 维仿射空间中对于一个线性变换，任意向量的象可以由它的基底的象来决定. 现在来看一看在欧氏空间中数积的概念与基底有怎样的关系？

设 $\boldsymbol{e}_1,\ \boldsymbol{e}_2,\ \cdots,\ \boldsymbol{e}_n$ 是 n 维仿射空间 $\mathfrak{R}$ 的一组基底，$\boldsymbol{x}$ 与 $\boldsymbol{y}$ 是 $\mathfrak{R}$ 的任二个向量，

$$\boldsymbol{x}=x_1\boldsymbol{e}_1+\cdots+x_n\boldsymbol{e}_n,$$

$$\boldsymbol{y}=y_1\boldsymbol{e}_1+\cdots+y_n\boldsymbol{e}_n.$$

首先，假定 $\mathfrak{R}$ 是欧氏空间，也就是说它里面任二个向量都有数积这个运算. 据定义，有

$$(\boldsymbol{x}, \boldsymbol{y})=\left(\sum_{i=1}^{n} x_i\boldsymbol{e}_i, \sum_{i=1}^{n} y_i\boldsymbol{e}_i\right)=\sum_{i,j=1}^{n} a_{ij}x_iy_j,$$

式中 $a_{ij}=(\boldsymbol{e}_i, \boldsymbol{e}_j)$，因之 $a_{ij}=a_{ji}$，即矩阵 $\boldsymbol{A}=(a_{ij})$ 是对称的. 当 $\boldsymbol{x}\neq\boldsymbol{0}$ 时，由于 $(\boldsymbol{x}, \boldsymbol{x})=\sum_{i,j=1}^{n} a_{ij}x_ix_j>0$，得知 $(\boldsymbol{x}, \boldsymbol{x})=\sum_{i,j=1}^{n} a_{ij}x_ix_j$ 是正定二次齐式.

反之，设 $\sum_{i,j=1}^{n} a_{ij}x_ix_j=\boldsymbol{X}'\boldsymbol{A}\boldsymbol{X}$ 是一个事先给定的正定二次齐式($a_{ij}=a_{ji}$). 如果这时用实数 $\sum_{i,j=1}^{n} a_{ij}x_iy_j$ 来定义向量 $\boldsymbol{x}=\sum x_i\boldsymbol{e}_i$ 与 $\boldsymbol{y}=\sum y_i\boldsymbol{e}_i$ 之数积，即定义 $(\boldsymbol{x}, \boldsymbol{y})=\sum_{i,j=1}^{n} a_{ij}x_iy_j$，那末容易验证公理 1°～4° 都是满足的. 例如

$$\begin{aligned}(\boldsymbol{y}, \boldsymbol{x})&=\sum_{i,j=1}^{n} a_{ij}y_ix_j=\sum_{i,j=1}^{n} a_{ji}y_ix_j \quad (\because\ a_{ij}=a_{ji})\\&=\sum_{i,j=1}^{n} a_{ji}x_jy_i=\sum_{i,j=1}^{n} a_{ij}x_iy_j=(\boldsymbol{x}, \boldsymbol{y}).\end{aligned}$$

并且，这时又显然有 $(\boldsymbol{e}_i, \boldsymbol{e}_j)=a_{ij}$.

总上所述，即证得了：

定理 1 使 n 维空间 $\mathfrak{R}$ 成为欧氏空间的充要条件是：由 $\mathfrak{R}$ 之任一组基底 $\boldsymbol{e}_1, \cdots, \boldsymbol{e}_n$ 中向量两两之数积所决定的 n 级矩阵

$$\begin{pmatrix}(\boldsymbol{e}_1, \boldsymbol{e}_1) & (\boldsymbol{e}_1, \boldsymbol{e}_2) & \cdots & (\boldsymbol{e}_1, \boldsymbol{e}_n)\\(\boldsymbol{e}_2, \boldsymbol{e}_1) & (\boldsymbol{e}_2, \boldsymbol{e}_2) & \cdots & (\boldsymbol{e}_2, \boldsymbol{e}_n)\\\vdots & \vdots & & \vdots\\(\boldsymbol{e}_n, \boldsymbol{e}_1) & (\boldsymbol{e}_n, \boldsymbol{e}_2) & \cdots & (\boldsymbol{e}_n, \boldsymbol{e}_n)\end{pmatrix}$$

为正定(实)对称矩阵.

简单地说，$\mathfrak{R}$ 成为欧氏空间的充要条件是由一个正定二次齐式所决定.

推论 n 维欧氏空间 $\mathfrak{R}$ 中任意 k 个($k\leqslant n$) 向量 $\boldsymbol{a}_1, \cdots,$

$\boldsymbol{\alpha}_k$ 所组成的 k 级行列式

$$G(\boldsymbol{\alpha}_1, \cdots, \boldsymbol{\alpha}_k)=\begin{vmatrix}(\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_1) & (\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2) & \cdots & (\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_k)\\ (\boldsymbol{\alpha}_2, \boldsymbol{\alpha}_1) & (\boldsymbol{\alpha}_2, \boldsymbol{\alpha}_2) & \cdots & (\boldsymbol{\alpha}_2, \boldsymbol{\alpha}_k)\\ \vdots & \vdots & & \vdots\\ (\boldsymbol{\alpha}_k, \boldsymbol{\alpha}_1) & (\boldsymbol{\alpha}_k, \boldsymbol{\alpha}_2) & \cdots & (\boldsymbol{\alpha}_k, \boldsymbol{\alpha}_k)\end{vmatrix}$$

不等于零的充要条件是 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_k$ 为线性无关的；并当 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_k$ 线性无关时，$G(\boldsymbol{\alpha}_1, \cdots, \boldsymbol{\alpha}_k)>0$†.

证明 当 $\boldsymbol{\alpha}_1, \cdots, \boldsymbol{\alpha}_k$ 线性无关时，可另取 $n-k$ 个向量 $\boldsymbol{\alpha}_{k+1}, \cdots, \boldsymbol{\alpha}_n$ 使 $\boldsymbol{\alpha}_1, \cdots, \boldsymbol{\alpha}_n$ 为欧氏空间 $\Re$ 的基底，于是，由定理 1，知

$$\begin{pmatrix}(\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_1) & \cdots & (\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_n)\\ \vdots & & \vdots\\ (\boldsymbol{\alpha}_n, \boldsymbol{\alpha}_1) & \cdots & (\boldsymbol{\alpha}_n, \boldsymbol{\alpha}_n)\end{pmatrix}$$

是正定的，因之 $G(\boldsymbol{\alpha}_1, \cdots, \boldsymbol{\alpha}_k)>0$.

反之，设 $G(\boldsymbol{\alpha}_1, \cdots, \boldsymbol{\alpha}_k)\neq 0$. 如果这时 $\boldsymbol{\alpha}_1, \cdots, \boldsymbol{\alpha}_k$ 线性相关，则可把 $\boldsymbol{\alpha}_k$（例如）写成 $\boldsymbol{\alpha}_1, \cdots, \boldsymbol{\alpha}_{k-1}$ 的线性组合：$\boldsymbol{\alpha}_k=\lambda_1\boldsymbol{\alpha}_1+\cdots+\lambda_{k-1}\boldsymbol{\alpha}_{k-1}$，因之，

$$G(\boldsymbol{\alpha}_1, \cdots, \boldsymbol{\alpha}_k)=\begin{vmatrix}(\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_1) & \cdots & (\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_{k-1}) & \left(\boldsymbol{\alpha}_1, \sum_1^{k-1}\lambda_i\boldsymbol{\alpha}_i\right)\\ \cdots & \cdots & \cdots & \cdots\\ (\boldsymbol{\alpha}_k, \boldsymbol{\alpha}_1) & \cdots & (\boldsymbol{\alpha}_k, \boldsymbol{\alpha}_{k-1}) & \left(\boldsymbol{\alpha}_k, \sum_1^{k-1}\lambda_i\boldsymbol{\alpha}_i\right)\end{vmatrix}$$

$$=\sum_{i=1}^{k-1}\lambda_i\cdot\begin{vmatrix}(\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_1)\cdots(\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_{k-1}) & (\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_i)\\ \cdots\cdots & \cdots\\ (\boldsymbol{\alpha}_k, \boldsymbol{\alpha}_1)\cdots(\boldsymbol{\alpha}_k, \boldsymbol{\alpha}_{k-1}) & (\boldsymbol{\alpha}_k, \boldsymbol{\alpha}_i)\end{vmatrix}=0,$$

与 $G(\boldsymbol{\alpha}_1, \cdots, \boldsymbol{\alpha}_k)\neq 0$ 之假定矛盾. 所以 $\boldsymbol{\alpha}_1, \cdots, \boldsymbol{\alpha}_k$ 一定是线

† 叫 $G(\boldsymbol{\alpha}_1, \cdots, \boldsymbol{\alpha}_k)$ 为 k 个向量 $\boldsymbol{\alpha}_1, \cdots, \boldsymbol{\alpha}_k$ 的格拉姆 (Gram) 行列式，而叫其相应之矩阵为 $\boldsymbol{\alpha}_1, \cdots, \boldsymbol{\alpha}_k$ 的格拉姆矩阵.

性无关的．证完．

格拉姆行列式的几何意义，留在下节讲了标准直交基底以后再谈．

从定理 1，又提出这样一个问题：当 $\boldsymbol{e}_1, \cdots, \boldsymbol{e}_n$ 和 $\boldsymbol{f}_1, \cdots, \boldsymbol{f}_n$ 是 n 维欧氏空间 $\mathfrak{R}$ 的两组基底时，它们对应的正定对称矩阵

$$\boldsymbol{S}_e=\begin{pmatrix}(\boldsymbol{e}_1, \boldsymbol{e}_1) & \cdots & (\boldsymbol{e}_1, \boldsymbol{e}_n)\\ \vdots & & \vdots\\ (\boldsymbol{e}_n, \boldsymbol{e}_1) & \cdots & (\boldsymbol{e}_n, \boldsymbol{e}_n)\end{pmatrix},$$

$$\boldsymbol{S}_f=\begin{pmatrix}(\boldsymbol{f}_1, \boldsymbol{f}_1) & \cdots & (\boldsymbol{f}_1, \boldsymbol{f}_n)\\ \vdots & & \vdots\\ (\boldsymbol{f}_n, \boldsymbol{f}_1) & \cdots & (\boldsymbol{f}_n, \boldsymbol{f}_n)\end{pmatrix}$$

之关系怎样？

事实上，若令基底变换的矩阵为 $\boldsymbol{Q}$，即 $\begin{pmatrix}\boldsymbol{f}_1\\ \vdots\\ \boldsymbol{f}_n\end{pmatrix}=\boldsymbol{Q}\cdot\begin{pmatrix}\boldsymbol{e}_1\\ \vdots\\ \boldsymbol{e}_n\end{pmatrix}$，则 $\boldsymbol{Q}$ 是满秩的；如再令 $\boldsymbol{x}=\sum x_i\boldsymbol{e}_i=\sum y_i\boldsymbol{f}_i$，则

$$(x_1, x_2, \cdots, x_n)=(y_1, y_2, \cdots, y_n)\cdot\boldsymbol{Q},$$

于是，

$$(\boldsymbol{x}, \boldsymbol{x})=(y_1, y_2, \cdots, y_n)\cdot\boldsymbol{S}_f\cdot\begin{pmatrix}y_1\\ y_2\\ \vdots\\ y_n\end{pmatrix}$$

$$=(x_1, x_2, \cdots, x_n)\cdot\boldsymbol{S}_e\cdot\begin{pmatrix}x_1\\ x_2\\ \vdots\\ x_n\end{pmatrix}$$

$$=(y_1, y_2, \cdots, y_n)\cdot QS_eQ'\cdot\begin{pmatrix}y_1\\y_2\\\vdots\\y_n\end{pmatrix},$$

因此，$QS_eQ'=S_f$. 这就证明了

定理 2 设欧氏空间 $\mathfrak{R}$ 中两组基底 $\boldsymbol{e}_1, \boldsymbol{e}_2, \cdots, \boldsymbol{e}_n$ 与 $\boldsymbol{f}_1, \boldsymbol{f}_2, \cdots, \boldsymbol{f}_n$ 所决定的两个正定对称矩阵(格拉姆矩阵)为 S_e 与 S_f，则 $S_f=QS_eQ'$，这里 Q 为基底变换的矩阵，即

$$\begin{pmatrix}\boldsymbol{f}_1\\\vdots\\\boldsymbol{f}_n\end{pmatrix}=Q\begin{pmatrix}\boldsymbol{e}_1\\\vdots\\\boldsymbol{e}_n\end{pmatrix}.$$

练 习 三 十

1. 设 $A=(a_{ij})$是正定对称矩阵，证明：不论 $x_1, \cdots, x_n, y_1, \cdots, y_n$ 是任何实数，恒有不等式

$$(\sum a_{ij}x_iy_j)^2\leqslant(\sum a_{ij}x_ix_j)(\sum a_{ij}y_iy_j).$$

2. 直接证明欧氏空间中 k 个向量 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_k$ 线性相关的充要条件是

$$\begin{vmatrix}(\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_1) & \cdots & (\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_k)\\\vdots & & \vdots\\(\boldsymbol{\alpha}_k, \boldsymbol{\alpha}_1) & \cdots & (\boldsymbol{\alpha}_k, \boldsymbol{\alpha}_k)\end{vmatrix}=0.$$

3. 在由 n 个实数组而成的空间 $\mathfrak{R}_n$ 中，按例 1 的方法使之成为欧氏空间. 今设 $P=(x_1, x_2, \cdots, x_n)$，$Q=(y_1, y_2, \cdots, y_n)$ 为 $\mathfrak{R}_n$ 的两点. 若点 M 在线段 PQ 上，且有 $\|\overrightarrow{PM}\|=\|\overrightarrow{QM}\|$，则叫 M 为线段 PQ 之中点. 试求：

i) M 点的坐标；

ii) 与 P, Q 等距离的点之轨迹叫超平面，求这超平面的方程；

iii) 又若 R, S 为这超平面上之二点，则 $(\overrightarrow{PQ}, \overrightarrow{RS})=0$，即 $PQ\perp RS$.

第 2 节　标准直交基底

我们知道：在平面解析几何里，选取了笛卡儿直角坐标系后，则分别在 x-轴与 y-轴上的基本单位向量 $\boldsymbol{i}$、$\boldsymbol{j}$ 可充当二

维空间 $\mathfrak{R}_2$ 的一组基底，且它们显然互相垂直而长都等于1. 同样，在 $\mathfrak{R}_3$ 中笛卡儿直角坐标系的 x-轴，y-轴、z-轴上的基本单位向量 $\boldsymbol{i}$、$\boldsymbol{j}$、$\boldsymbol{k}$ 是 $\mathfrak{R}_3$ 之一组基底，且它们两两互相垂直而长都等于1. 又，由前节例1得知：给了仿射空间 $\mathfrak{R}$ 的一组基底 $\boldsymbol{e}_1, \cdots, \boldsymbol{e}_n$，若 $\boldsymbol{x}=\sum\limits_1^n x_i\boldsymbol{e}_i$，$\boldsymbol{y}=\sum\limits_1^n y_i\boldsymbol{e}_i$，当规定 $(\boldsymbol{x}, \boldsymbol{y})=\sum\limits_1^n x_iy_i$ 后，$\mathfrak{R}$ 就成为欧氏空间. 这时，由基底 $\boldsymbol{e}_1, \cdots, \boldsymbol{e}_n$ 所确定的正定对称矩阵是单位矩阵，也就是说 $(\boldsymbol{e}_i, \boldsymbol{e}_j)=\delta_{ij}$($\delta_{ij}=0$ 当 $i\neq j$ 时，$\delta_{ii}=1$)，即 $\boldsymbol{e}_1, \cdots, \boldsymbol{e}_n$ 中两两互相直交且每 $\boldsymbol{e}_i$ 的长等于1. 欧氏空间 $\mathfrak{R}$ 中象这样的一组基底叫做*标准直交基底*，即

定义 假定 $\boldsymbol{e}_1, \cdots, \boldsymbol{e}_n$ 是 n 维欧氏空间 $\mathfrak{R}$ 的基底. 如果它们两两直交，我们就说它们是*直交基底*；若除了两两垂直以外，它们的长又都等于1，即

$$(\boldsymbol{e}_i, \boldsymbol{e}_j)=\begin{cases}1, & \text{当 } i=j \text{ 时;}\\ 0, & \text{当 } i\neq j \text{ 时;}\end{cases}$$

我们就说它们是*标准直交基底*.

于是由直交基底 $\boldsymbol{e}_1, \cdots, \boldsymbol{e}_n$，只要把 $\boldsymbol{e}_i$ 标准化令 $\boldsymbol{e}'_i=\dfrac{\boldsymbol{e}_i}{\|\boldsymbol{e}_i\|}$，那末 $\boldsymbol{e}'_1, \cdots, \boldsymbol{e}'_n$ 就是标准直交基底了.

当空间 $\mathfrak{R}$ 是三维(或二维)时，标准直交基底的意义与上面说的 $\boldsymbol{i}, \boldsymbol{j}, \boldsymbol{k}$ (或 $\boldsymbol{i}, \boldsymbol{j}$)之意义相同，因之标准直交基底在欧氏空间中的作用就象解析几何里而直角坐标系一样，任一个向量 $\boldsymbol{x}=\sum x_i\boldsymbol{e}_i$ 在标准直交基底 $\boldsymbol{e}_1, \cdots, \boldsymbol{e}_n$ 内的坐标 $x_1, \cdots, x_n$ 也相当于解析几何里在直角坐标系内点的坐标. 由是，在欧氏空间里，我们总是采用标准直交基底.

在欧氏空间中，欲求标准直交基底，只要先求出直交基底

即可．今问欧氏空间是否有直交基底？读者可能会发生误解，认为这是一个不成问题的问题．因为在前节的例1就说过，不管 $\boldsymbol{e}_1, \cdots, \boldsymbol{e}_n$ 是 n 维仿射空间 $\mathfrak{R}$ 之任何基底，如果将二个向量 $\boldsymbol{x}=\sum\limits_1^n \xi_i\boldsymbol{e}_i$, $\boldsymbol{y}=\sum\limits_1^n \eta_i\boldsymbol{e}_i$ 的数积定义为 $(\boldsymbol{x}, \boldsymbol{y})=\sum\limits_1^n \xi_k\eta_k$ 使 $\mathfrak{R}$ 变成欧氏空间后，则 $\boldsymbol{e}_1, \cdots, \boldsymbol{e}_n$ 就是 $\mathfrak{R}$ 的标准直交基底了．但要注意，这里对于给定的一组基底 $\boldsymbol{e}_1, \cdots, \boldsymbol{e}_n$，我们是采用特殊的正定对称矩阵，即单位矩阵，来定义它们的数积（参照第1节定理1）．可是，在 n 维欧氏空间 $\mathfrak{R}$ 中，对于一组基底 $\boldsymbol{e}_1, \cdots, \boldsymbol{e}_n$，我们只能有 $(\boldsymbol{e}_i, \boldsymbol{e}_j)=a_{ij}$, $a_{ij}=a_{ji}$，而 $\boldsymbol{A}=(a_{ij})$ 是正定的；就是说，是先令 $(\boldsymbol{e}_i, \boldsymbol{e}_j)=a_{ij}$ 而 $\boldsymbol{A}=(a_{ij})$ 为事先给定的一个正定对称矩阵，才使仿射空间 $\mathfrak{R}$ 变为欧氏空间的．于是，这时 $\boldsymbol{e}_1, \cdots, \boldsymbol{e}_n$ 就不一定为 $\mathfrak{R}$ 的直交基底了，因之我们的问题也就在这个地方，即是说：这时虽然 $(\boldsymbol{e}_i, \boldsymbol{e}_j)=a_{ij}$，但在 $\mathfrak{R}$ 中能否找着另一组基底 $\boldsymbol{f}_1, \cdots, \boldsymbol{f}_n$ 使它们两两直交呢？所以说，要解决直交基底或标准直交基底的存在性，是欧氏空间里的一个重要问题．我们又知道，任何一组基底中的向量必是线性无关的，于是要解决直交基底的存在性，首先就要解决这样一个问题，即一组两两直交的非零向量是否必为线性无关的呢？如果说两两直交的向量组必为线性相关，那末当然不会有直交基底存在．

定理1 若 $\boldsymbol{e}_1, \boldsymbol{e}_2, \cdots, \boldsymbol{e}_k$ 是一组两两直交的非零向量，那末它们一定线性无关．

证明 设 $\lambda_1\boldsymbol{e}_1+\lambda_2\boldsymbol{e}_2+\cdots+\lambda_k\boldsymbol{e}_k=\boldsymbol{0}$．于是因 $(\boldsymbol{e}_i, \boldsymbol{e}_j)=0$ $(i\neq j)$ 及 $(\boldsymbol{e}_i, \boldsymbol{e}_i)\neq 0$，就得到

$$\begin{aligned}0&=(\lambda_1\boldsymbol{e}_1+\lambda_2\boldsymbol{e}_2+\cdots+\lambda_k\boldsymbol{e}_k, \boldsymbol{e}_i)\\&=\sum_{t=1}^{k}\lambda_t\cdot(\boldsymbol{e}_t, \boldsymbol{e}_i)=\lambda_i\cdot(\boldsymbol{e}_i, \boldsymbol{e}_i),\end{aligned}$$

故得 $\lambda_i=0$, 即 $\boldsymbol{e}_1, \cdots, \boldsymbol{e}_k$ 线性无关. 证完.

于是, 在 n 维欧氏空间 $\mathfrak{R}$ 中, 两两直交的非零向量之个数不能多于 n, 任意 n 个两两直交非零向量必为 $\mathfrak{R}$ 的一组基底(直交基底). 下面来证明 n 维欧氏空间 $\mathfrak{R}$ 确有 n 个两两直交的非零向量的组, 即:

定理 2 *n 维欧氏空间 $\mathfrak{R}$ 必有直交基底, 因之也有标准直交基底.*

证明 设 $\boldsymbol{e}_1, \cdots, \boldsymbol{e}_n$ 为 $\mathfrak{R}$ 之一组基底, 于是由第 1 节定理 1, 即知矩阵

$$\boldsymbol{A}=\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} \quad (\text{但 } a_{ij}=(\boldsymbol{e}_i, \boldsymbol{e}_j))$$

为正定(实)对称矩阵, 故有满秩(实)矩阵 $\boldsymbol{Q}$ 使 $\boldsymbol{QAQ'}=\boldsymbol{E}$, 因而令

$$\begin{pmatrix} \boldsymbol{f}_1 \\ \boldsymbol{f}_2 \\ \vdots \\ \boldsymbol{f}_n \end{pmatrix}=\boldsymbol{Q}\cdot\begin{pmatrix} \boldsymbol{e}_1 \\ \boldsymbol{e}_2 \\ \vdots \\ \boldsymbol{e}_n \end{pmatrix},$$

则 $\boldsymbol{f}_1, \cdots, \boldsymbol{f}_n$ 亦为 $\mathfrak{R}$ 之一组基底(因 $\boldsymbol{Q}$ 满秩), 且由第 1 节定理 2, 知 $\boldsymbol{f}_1, \cdots, \boldsymbol{f}_n$ 为 $\mathfrak{R}$ 之一组标准直交基底†.

上述的证法简单, 而且还联系了二次齐式, 即揭示了六、七两章的内在连系, 同时一下子解决了标准直交基底的存在性, 这是这个方法的优点. 但这个方法只说了标准直交基底

---

† 或可直接证明于下: 因 $\boldsymbol{f}_i=q_{i1}\boldsymbol{e}_1+\cdots+q_{in}\boldsymbol{e}_n\ (i=1, \cdots, n)$, 故 $(\boldsymbol{f}_i, \boldsymbol{f}_j)=\left(\sum_{s=1}^{n} q_{is}\boldsymbol{e}_s, \sum_{t=1}^{n} q_{jt}\boldsymbol{e}_t\right)=\sum_{s,t=1}^{n} q_{is}q_{jt}\cdot(\boldsymbol{e}_s, \boldsymbol{e}_t)=\sum_{s,t=1}^{n} q_{is}a_{st}q_{jt}$, 后者显然是矩阵 $\boldsymbol{QAQ'}(=\boldsymbol{E})$ 的第 i 行第 j 列处的元素, 即等于 δ_{ij}. 故 $\boldsymbol{f}_1, \cdots, \boldsymbol{f}_n$ 为 $\mathfrak{R}$ 之标准直交基底.

的存在，未涉及它的实际求法，这是此方法的缺点．下面再讲一个方法，这方法虽较烦，但它的优点是构造性的，即从任何一组基底 $\boldsymbol{e}_1, \cdots, \boldsymbol{e}_n$ 出发，可以有具体的方法找得一组直交基底 $\boldsymbol{f}_1, \cdots, \boldsymbol{f}_n$.

先从直觉空间谈起，从中看出思想方法．例如，在三维空间 $\mathfrak{R}_3$ 中，有三个向量 $\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3$ 不共面，如图 7-1．于是，$\mathfrak{R}_3$ 中任何向量都是 $\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3$ 的线性组合，即 $\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3$ 为基底．所谓求直交基底，就是要通过 O 点作三个两两互相垂直的向量，最简单的方法就是先作含向量 $\boldsymbol{e}_1, \boldsymbol{e}_2$ 的平面 π，再过 O 点作垂直于平面 π 的向量 $\boldsymbol{f}_3$，然后在平面 π 内过 O 点作向量 $\boldsymbol{f}_2 \perp \boldsymbol{e}_1$，那末三个向量 $\boldsymbol{e}_1, \boldsymbol{f}_2, \boldsymbol{f}_3$ 就两两直交了．因此，我们先令 $\boldsymbol{f}_1 = \boldsymbol{e}_1$; 再在含 $\boldsymbol{e}_1 = \boldsymbol{f}_1$ 与 $\boldsymbol{e}_2$ 的平面 π 上过 O 点作 $\boldsymbol{f}_2 \perp \boldsymbol{f}_1$，因为 $\boldsymbol{f}_2$ 在 π 上，所以 $\boldsymbol{f}_2$ 可写为 $\boldsymbol{e}_2$ 与 $\boldsymbol{e}_1 = \boldsymbol{f}_1$ 的线性组合，又因为 $\boldsymbol{f}_2$ 的长可以任意，所以在线性组合中可以令 $\boldsymbol{e}_2$ 的系数等于 1，即 $\boldsymbol{f}_2 = \boldsymbol{e}_2 + \alpha \boldsymbol{f}_1$. 最后，因 $\boldsymbol{f}_1, \boldsymbol{f}_2, \boldsymbol{e}_3$ 不共面，而 $\boldsymbol{f}_3$ 的长又可任意，所以 $\boldsymbol{f}_3$ 得写为 $\boldsymbol{f}_3 = \boldsymbol{e}_3 + \lambda \boldsymbol{f}_1 + \mu \boldsymbol{f}_2$.

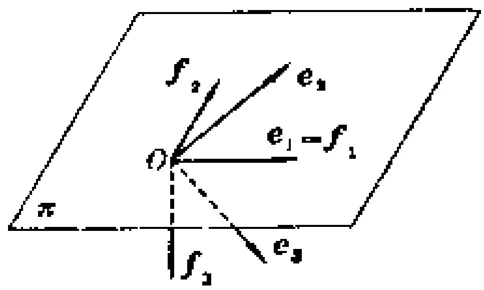


图 7-1

现在就根据这个思想，来从 $\boldsymbol{e}_1, \cdots, \boldsymbol{e}_n$ 去找直交基底 $\boldsymbol{f}_1, \cdots, \boldsymbol{f}_n$.

首先，就令 $\boldsymbol{f}_1 = \boldsymbol{e}_1$，故 $\boldsymbol{f}_1 \neq \boldsymbol{0}$，再来逐步地求 $\boldsymbol{f}_2, \cdots, \boldsymbol{f}_n$ 使得 $(\boldsymbol{f}_i, \boldsymbol{f}_j) = 0 (i \neq j)$，具体的步骤如下：

作向量 $\boldsymbol{f}_2 = \boldsymbol{e}_2 + \alpha \boldsymbol{f}_1$，$\alpha$ 是待定的实数．由 $(\boldsymbol{f}_2, \boldsymbol{f}_1) = (\boldsymbol{e}_2, \boldsymbol{f}_1) + \alpha \cdot (\boldsymbol{f}_1, \boldsymbol{f}_1) = 0$，得 $\alpha = -\dfrac{(\boldsymbol{e}_2, \boldsymbol{f}_1)}{(\boldsymbol{f}_1, \boldsymbol{f}_1)}$，这时，因为 $\boldsymbol{e}_1$、$\boldsymbol{e}_2$ 线性无关，所以 $\boldsymbol{f}_2 \neq \boldsymbol{0}$.

再用归纳法，假定两两垂直而又都不为零的向量 $\boldsymbol{f}_1$, $\boldsymbol{f}_2$, $\cdots$, $\boldsymbol{f}_{k-1}$ 已经作出，即 $(\boldsymbol{f}_i, \boldsymbol{f}_j)=0(i\neq j;\ i,\ j=1,\ \cdots,\ n)$，且 $\boldsymbol{f}_i$ 为 $\boldsymbol{e}_1$, $\cdots$, $\boldsymbol{e}_i$ 的线性组合．再令

$$\boldsymbol{f}_k=\boldsymbol{e}_k+\lambda_1\boldsymbol{f}_1+\lambda_2\boldsymbol{f}_2+\cdots+\lambda_{k-1}\boldsymbol{f}_{k-1},$$

这里 λ_1, λ_2, $\cdots$, λ_{k-1} 是待定的实数；如果 $(\boldsymbol{f}_k, \boldsymbol{f}_i)=0$, $(i=1, 2, \cdots, k-1)$，则 $\boldsymbol{f}_1$, $\boldsymbol{f}_2$, $\cdots$, $\boldsymbol{f}_{k-1}$, $\boldsymbol{f}_k$ 就两两直交，且又都非零向量；但由

$$\begin{aligned}0=(\boldsymbol{f}_k,\ \boldsymbol{f}_i)&=(\boldsymbol{e}_k,\ \boldsymbol{f}_i)+\sum_{t=1}^{k-1}\lambda_t\cdot(\boldsymbol{f}_t,\ \boldsymbol{f}_i)\\&=(\boldsymbol{e}_k,\ \boldsymbol{f}_i)+\lambda_i\cdot(\boldsymbol{f}_i,\ \boldsymbol{f}_i),\quad (i=1,\ \cdots,\ k-1)\end{aligned}$$

得

$$\lambda_i=-\frac{(\boldsymbol{e}_k,\ \boldsymbol{f}_i)}{(\boldsymbol{f}_i,\ \boldsymbol{f}_i)},\ (i=1,\ \cdots,\ k-1),$$

即确可求出 $\boldsymbol{f}_k$．此即表明：从 $\boldsymbol{e}_1$, $\cdots$, $\boldsymbol{e}_{k-1}$ 出发若能求得出两两直交的 $\boldsymbol{f}_1$, $\cdots$, $\boldsymbol{f}_{k-1}$, 那末从 $\boldsymbol{e}_1$, $\cdots$, $\boldsymbol{e}_{k-1}$, $\boldsymbol{e}_k$ 也就能求得出两两直交的 $\boldsymbol{f}_1$, $\cdots$, $\boldsymbol{f}_{k-1}$, $\boldsymbol{f}_k$.

继续这样做，一直把所给出的 $\boldsymbol{e}_1$, $\cdots$, $\boldsymbol{e}_n$ 都用完，我们就得到 n 个非零且又两两直交的向量 $\boldsymbol{f}_1$, $\boldsymbol{f}_2$, $\cdots$, $\boldsymbol{f}_n$, 于是 $\boldsymbol{f}_1$, $\boldsymbol{f}_2$, $\cdots$, $\boldsymbol{f}_n$ 是 $\mathfrak{R}$ 的直交基底．再把 $\boldsymbol{f}_i$ 标准化，令 $\boldsymbol{e}'_i=\boldsymbol{f}_i/\|\boldsymbol{f}_i\|$, $(i=1, 2, \cdots, n)$，则 $\boldsymbol{e}'_1$, $\cdots$, $\boldsymbol{e}'_n$ 就是 $\mathfrak{R}$ 的标准直交基底，证完．

上面所讲的直交化过程，不只是告诉了我们从任一组基底怎样去求直交基底的方法，同时还告诉了我们从 k 个线性无关的向量怎样去求 k 个两两直交的向量．

[例]　设 $2n+1$ 个连续函数

$$1,\ \cos t,\ \sin t,\ \cos 2t,\ \sin 2t,\ \cdots,\ \cos nt,\ \sin nt$$

的定义域都是区间 $[0,\ 2\pi]$，它们的线性组合

$$P(t)=a_0+a_1\cos t+b_1\sin t+\cdots+a_n\cos nt+b_n\sin nt$$

叫做 n 次三角多项式. 显然，所有这样的 n 次三角多项式形成一个 $2n+1$ 维的空间 $\mathfrak{R}$，叫做 n 次三角多项式空间. 在 $\mathfrak{R}$ 内，两个 n 次三角多项式 $P(t)$、$Q(t)$ 的数积，可以由积分

$$(P(t),\ Q(t))=\int_0^{2\pi} P(t)\cdot Q(t)\,dt$$

来定义，这样，$\mathfrak{R}$ 就成为欧氏空间了. 因 $\int_0^{2\pi}\cos kt\cos lt\,dt=0$ $(k\neq l)$，$\int_0^{2\pi}\sin kt\cdot\sin lt\,dt=0\,(k\neq l)$，$\int_0^{2\pi}\sin\lambda t\cdot\cos\mu t\,dt=0$，$\int_0^{2\pi}\cos kt\,dt=0$，$\int_0^{2\pi}\sin kt\,dt=0$，所以 1，$\cos t$，$\sin t$，$\cdots$，$\cos nt$、$\sin nt$ 为 $\mathfrak{R}$ 之直交基底. 又因为

$$\int_0^{2\pi}\sin^2 kt\cdot dt=\int_0^{2\pi}\cos^2 kt\cdot dt=\pi,\quad \int_0^{2\pi} 1\,dt=2\pi,$$

故 $\frac{1}{\sqrt{2\pi}}$，$\frac{1}{\sqrt{\pi}}\cos t$，$\frac{1}{\sqrt{\pi}}\sin t$，$\cdots$，$\frac{1}{\sqrt{\pi}}\cos nt$，$\frac{1}{\sqrt{\pi}}\sin nt$ 为 $\mathfrak{R}$ 之标准直交基底.

今后恒假定 $\mathfrak{R}$ 是 n 维欧氏空间，而

$$\boldsymbol{e}_1,\ \boldsymbol{e}_2,\ \cdots,\ \boldsymbol{e}_n \tag{5}$$

是 $\mathfrak{R}$ 的一组标准直交基底. 现在来研究下面几个问题.

第一个问题：$\mathfrak{R}$ 中向量 $\boldsymbol{x}$ 在标准直交基底(5)内的坐标等于什么？

设 $\boldsymbol{x}=\xi_1\boldsymbol{e}_1+\xi_2\boldsymbol{e}_2+\cdots+\xi_n\boldsymbol{e}_n$，则

$$(\boldsymbol{x},\ \boldsymbol{e}_i)=\sum_{t=1}^{n}\xi_t\cdot(\boldsymbol{e}_t,\ \boldsymbol{e}_i)=\xi_i,$$

这就是说：向量在标准直交基底内的坐标等于这向量与各个对应的基底向量之数积.

上面这个论断是有几何意义的. 我们知道，在解析几何里，向量 $\boldsymbol{x}$ 与单位向量 $\boldsymbol{e}$ 的数积 $(\boldsymbol{x},\ \boldsymbol{e})$ 等于 $\boldsymbol{x}$ 在 $\boldsymbol{e}$ 上的射

影，这是因为：

$$\cos(\widehat{x, e})=\frac{(x, e)}{\|x\|\cdot\|e\|}=\frac{(x, e)}{\|x\|},$$

所以 $(x, e)=\|x\|\cdot\cos(\widehat{x, e})=x$ 在 e 上的射影．因此，也叫 ξ_i 是 x 在 e_i 上的射影：$\xi_i=(x, e_i)=\|x\|\cdot\cos(x, e_i)$，也就是说，向量在标准直交基底内之坐标等于它在基底向量上的射影．它的逆定理也成立（留给读者当练习题去作，参看练习三十一的第 1 题）．

第二个问题：$\Re$ 中二向量的数积在标准直交基底 (5) 内的表达式怎样？

设 $x=\sum\limits_{i=1}^{n}\xi_i e_i,\ y=\sum\limits_{j=1}^{n}\mu_j e_j$，则

$$\begin{aligned}(x, y)&=\left(\sum_{i=1}^{n}\xi_i e_i,\ \sum_{j=1}^{n}\mu_j e_j\right)\\&=\sum_{i,j=1}^{n}\xi_i\mu_j(e_i, e_j)=\sum_{i=1}^{n}\xi_i\mu_i,\end{aligned}$$

这是说：二向量之数积如用它们在标准直交基底内的坐标来表示，就等于这二个向量的相应坐标乘积之和．它的逆定理也成立（参看练习三十一的第 2 题）．这个性质实际上早就知道了，因为这无异乎是说由单位矩阵这个特殊的正定对称矩阵所定义的数积时，它的基底就是标准直交基底．

第三个问题：怎样借助标准直交基底说明格拉姆行列式的几何意义？

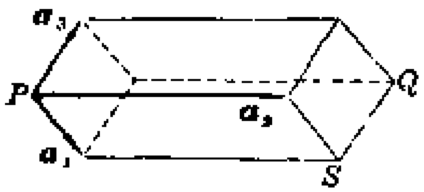


图 7-2

事实上，即令基底不是两两直交时，格拉姆行列式也是有几何意义的，我们先从直觉的三维空间 $\Re_3$ 来谈起．

如图 7-2，a_1，a_2，a_3 是从点 P 引出的三个不共面的向

量，可用坐标表示（在直角坐标系中）为 $\boldsymbol{a}_1=\{x_1, y_1, z_1\}$，$\boldsymbol{a}_2=\{x_2, y_2, z_2\}$，$\boldsymbol{a}_3=\{x_3, y_3, z_3\}$，于是三个向量 $\boldsymbol{a}_i(i=1, 2, 3)$ 的格拉姆行列式

$$G(\boldsymbol{a}_1, \boldsymbol{a}_2, \boldsymbol{a}_3)=\begin{vmatrix}(\boldsymbol{a}_1, \boldsymbol{a}_1) & (\boldsymbol{a}_1, \boldsymbol{a}_2) & (\boldsymbol{a}_1, \boldsymbol{a}_3)\\ (\boldsymbol{a}_2, \boldsymbol{a}_1) & (\boldsymbol{a}_2, \boldsymbol{a}_2) & (\boldsymbol{a}_2, \boldsymbol{a}_3)\\ (\boldsymbol{a}_3, \boldsymbol{a}_1) & (\boldsymbol{a}_3, \boldsymbol{a}_2) & (\boldsymbol{a}_3, \boldsymbol{a}_3)\end{vmatrix}$$

$$=\begin{vmatrix}\sum x_1^2 & \sum x_1x_2 & \sum x_1x_3\\ \sum x_1x_2 & \sum x_2^2 & \sum x_2x_3\\ \sum x_1x_3 & \sum x_2x_3 & \sum x_3^2\end{vmatrix}$$

$$=\begin{vmatrix}x_1 & y_1 & z_1\\ x_2 & y_2 & z_2\\ x_3 & y_3 & z_3\end{vmatrix}\cdot\begin{vmatrix}x_1 & x_2 & x_3\\ y_1 & y_2 & y_3\\ z_1 & z_2 & z_3\end{vmatrix}$$

$$=\begin{vmatrix}x_1 & y_1 & z_1\\ x_2 & y_2 & z_2\\ x_3 & y_3 & z_3\end{vmatrix}^2,$$

即平行六面体 $P—Q$ 之体积的平方；而

$$G(\boldsymbol{a}_1, \boldsymbol{a}_2)=\begin{vmatrix}(\boldsymbol{a}_1, \boldsymbol{a}_1) & (\boldsymbol{a}_1, \boldsymbol{a}_2)\\ (\boldsymbol{a}_2, \boldsymbol{a}_1) & (\boldsymbol{a}_2, \boldsymbol{a}_2)\end{vmatrix}$$

$$=(\boldsymbol{a}_1, \boldsymbol{a}_1)\cdot(\boldsymbol{a}_2, \boldsymbol{a}_2)-(\boldsymbol{a}_1, \boldsymbol{a}_2)^2$$

$$=\|\boldsymbol{a}_1\|^2\cdot\|\boldsymbol{a}_2\|^2-\|\boldsymbol{a}_1\|^2\cdot\|\boldsymbol{a}_2\|^2\cdot\cos^2(\widehat{\boldsymbol{a}_1, \boldsymbol{a}_2})$$

$$=\|\boldsymbol{a}_1\|^2\cdot\|\boldsymbol{a}_2\|^2\cdot\sin^2(\widehat{\boldsymbol{a}_1, \boldsymbol{a}_2}),$$

即平行四边形 $P—S$ 之面积的平方．

由于这个性质，我们就叫 n 维欧氏空间 $\mathfrak{R}$ 里面 k 个线性无关向量 $\boldsymbol{e}_1, \cdots, \boldsymbol{e}_n$ 的格拉姆行列式

$$G(\boldsymbol{e}_1, \boldsymbol{e}_2, \cdots, \boldsymbol{e}_k)=\begin{vmatrix}(\boldsymbol{e}_1, \boldsymbol{e}_1) & (\boldsymbol{e}_1, \boldsymbol{e}_2) & \cdots & (\boldsymbol{e}_1, \boldsymbol{e}_k)\\ (\boldsymbol{e}_2, \boldsymbol{e}_1) & (\boldsymbol{e}_2, \boldsymbol{e}_2) & \cdots & (\boldsymbol{e}_2, \boldsymbol{e}_k)\\ \vdots & \vdots & & \vdots\\ (\boldsymbol{e}_k, \boldsymbol{e}_1) & (\boldsymbol{e}_k, \boldsymbol{e}_2) & \cdots & (\boldsymbol{e}_k, \boldsymbol{e}_k)\end{vmatrix}$$

为 $\mathfrak{R}$ 里面 k 维超平行体之体积的平方，而 $\boldsymbol{e}_1, \cdots, \boldsymbol{e}_k$ 叫做这超平行体的棱向量.

因相应于 $G(\boldsymbol{e}_1, \cdots, \boldsymbol{e}_k)$ 之矩阵为正定的，故由练习二十九之第 11 题，可知

$$G(\boldsymbol{e}_1, \cdots, \boldsymbol{e}_k)\leqslant(\boldsymbol{e}_1, \boldsymbol{e}_1)\cdot(\boldsymbol{e}_2, \boldsymbol{e}_2)\cdots(\boldsymbol{e}_k, \boldsymbol{e}_k)$$
$$=|\boldsymbol{e}_1|^2\cdot|\boldsymbol{e}_2|^2\cdots|\boldsymbol{e}_k|^2,$$

且等号成立的充要条件是 $(\boldsymbol{e}_i, \boldsymbol{e}_j)=0(i\neq j)$. 这就是说，欧氏空间 $\mathfrak{R}$ 里面 k 维超平行体之体积不大于它的棱向量之长之乘积，而且在且仅在棱向量两两垂直时，这体积才等于棱向量之长的乘积. 这个结论恰好与直觉空间中的情况完全一致. 因此以标准直交基底为棱向量的 n 维超平行体之体积等于1，这也正是三维空间中边长为1的正立方体的体积等于1(或二维空间中边长为1的正方形的面积等于1) 这一个事实的推广. 正因为这样，由标准直交基底所组成的格拉姆行列式的平方根，叫做边长为1的 n 维超立方体的体积；而由直交基底所组成的格拉姆行列式之平方根，叫做 n 维超长方体的体积.

练习三十一

1. 设 $\boldsymbol{e}_1, \boldsymbol{e}_2, \cdots, \boldsymbol{e}_n$ 为欧氏空间 $\mathfrak{R}$ 的一组基底，如若任一向量 $\boldsymbol{x}=\sum\limits_{i=1}^{n}\xi_i\boldsymbol{e}_i$ 在这基底内之坐标 ξ_i 等于 $\boldsymbol{x}$ 在 $\boldsymbol{e}_i$ 上的射影 $|\boldsymbol{x}|\cdot\cos(\widehat{\boldsymbol{x}, \boldsymbol{e}_i})$，则 $\boldsymbol{e}_1, \boldsymbol{e}_2, \cdots, \boldsymbol{e}_n$ 是 $\mathfrak{R}$ 的标准直交基底.

2. 设 $\boldsymbol{e}_1, \boldsymbol{e}_2, \cdots, \boldsymbol{e}_n$ 是欧氏空间 $\mathfrak{R}$ 的一组基底. 若由 $\boldsymbol{x}=\sum\limits_{i=1}^{n}\xi_i\boldsymbol{e}_i$, $\boldsymbol{y}=$

$\sum_{i=1}^{n}\eta_i e_i$ 恒得 $(x, y)=\sum_{i=1}^{n}\xi_i\eta_i$，试由标准直交基底的定义直接证明 $e_1, \cdots, e_n$ 为 $\mathfrak{R}$ 之标准直交基底。

3. 设 $e_1, \cdots, e_n$ 为欧氏空间 $\mathfrak{R}$ 的标准直交基底；$x_i=\sum_{l=1}^{n}\xi_{il}e_l (i=1, \cdots, k)$ 是 $\mathfrak{R}$ 的任 k 个向量。试证：$x_1, \cdots, x_k$ 的格拉姆矩阵等于 $A\cdot A'$，这里

$$A=\begin{pmatrix}\xi_{11} & \xi_{12} & \cdots & \xi_{1n}\\ \xi_{21} & \xi_{22} & \cdots & \xi_{2n}\\ \vdots & \vdots & & \vdots\\ \xi_{k1} & \xi_{k2} & \cdots & \xi_{kk}\end{pmatrix}.$$

4. 设 $e_1, e_2, \cdots, e_n$ 是欧氏空间 $\mathfrak{R}$ 的标准直交基底；$x_1, x_2, \cdots, x_k$ 是 $\mathfrak{R}$ 的任 k 个向量。试证：$x_1, x_2, \cdots, x_k$ 两两直交的充要条件是

$$\sum_{l=1}^{n}(x_i, e_l)\cdot(x_j, e_l)=0, \quad (i, j=1, 2, \cdots, k; i\neq j).$$

5. 哈达马 (Hadamard) 不等式：设 $A=(a_{ij})$ 是任一个 n 级满秩矩阵，则 $|A|\cdot|\overline{A}|\leqslant\prod_{k=1}^{n}\{a_{1k}\overline{a}_{1k}+\cdots+a_{nk}\overline{a}_{nk}\}$。并问等号“=”成立的条件是什么？几何意义怎样？

第 3 节　直交变换与直交矩阵

在上一节里，我们知道欧氏空间 $\mathfrak{R}$ 有标准直交基底，而且还不止一组。今假定

$$e_1, e_2, \cdots, e_n \tag{6}$$

与

$$f_1, f_2, \cdots, f_n \tag{7}$$

是 $\mathfrak{R}$ 的两组标准直交基底，那末由基底(6)转化为基底(7)的满秩矩阵 Q 以及 $\mathfrak{R}$ 里面使 e_i 的象为 $f_i (i=1, \cdots, n)$ 的唯一线性变换 $\mathscr{T}$（即 $e_i\mathscr{T}=f_i$; $i=1, 2, \cdots, n$）有什么特征呢？本节就讨论这个问题。

首先，注意到

$$\begin{pmatrix} \boldsymbol{f}_1 \\ \vdots \\ \boldsymbol{f}_n \end{pmatrix} = \begin{pmatrix} \boldsymbol{e}_1\mathscr{T} \\ \vdots \\ \boldsymbol{e}_n\mathscr{T} \end{pmatrix} = \boldsymbol{Q}\cdot\begin{pmatrix} \boldsymbol{e}_1 \\ \vdots \\ \boldsymbol{e}_n \end{pmatrix}, \tag{8}$$

而因(6)与(7)都是$\mathfrak{R}$的基底，故$\mathfrak{R}$的任何向量$\boldsymbol{x}$都可写为$\boldsymbol{x}=\sum\xi_i\boldsymbol{e}_i=\sum\eta_i\boldsymbol{f}_i$形，这时

$$(\xi_1, \cdots, \xi_n)=(\eta_1, \cdots, \eta_n)\cdot\boldsymbol{Q}. \tag{9}$$

为了看出线性变换$\mathscr{T}$的特性，从(8)则得到

$$\boldsymbol{x}\mathscr{T}=(\sum\xi_i\boldsymbol{e}_i)\mathscr{T}=\sum\xi_i\cdot(\boldsymbol{e}_i\mathscr{T})=\sum\xi_i\boldsymbol{f}_i,$$

于是，由于$\boldsymbol{f}_1, \cdots, \boldsymbol{f}_n$是标准直交基底，得知$(\boldsymbol{x}\mathscr{T}, \boldsymbol{x}\mathscr{T})=\sum\xi_i^2$；另方面，从$\boldsymbol{x}=\sum\xi_i\boldsymbol{e}_i$并利用(6)之标准直交性，又得到$(\boldsymbol{x}, \boldsymbol{x})=\sum\xi_i^2$；于是$(\boldsymbol{x}\mathscr{T}, \boldsymbol{x}\mathscr{T})=(\boldsymbol{x}, \boldsymbol{x})$，说明了象(8)这样的线性变换$\mathscr{T}$使空间$\mathfrak{R}$中任意向量$\boldsymbol{x}$的长都不变，这样的变换在直觉空间中显然是刚体运动，而在欧氏空间中因为它是由标准直交基底所引起的，故而称为直交变换，即：

定义 1 使欧氏空间$\mathfrak{R}$中任意向量$\boldsymbol{x}$的长都保持不变的线性变换$\mathscr{T}$，即$(\boldsymbol{x}\mathscr{T}, \boldsymbol{x}\mathscr{T})=(\boldsymbol{x}, \boldsymbol{x})$，叫**直交变换**.

由定义 1，易知：

定理 1 线性变换$\mathscr{T}$为直交变换的充要条件是$\mathscr{T}$使任二个向量$\boldsymbol{x}$、$\boldsymbol{y}$的数积保持不变，即$(\boldsymbol{x}\mathscr{T}, \boldsymbol{y}\mathscr{T})=(\boldsymbol{x}, \boldsymbol{y})$.

证明 若对任二个向量$\boldsymbol{x}$, $\boldsymbol{y}$有$(\boldsymbol{x}\mathscr{T}, \boldsymbol{y}\mathscr{T})=(\boldsymbol{x}, \boldsymbol{y})$，特令$\boldsymbol{y}=\boldsymbol{x}$时就有$(\boldsymbol{x}\mathscr{T}, \boldsymbol{x}\mathscr{T})=(\boldsymbol{x}, \boldsymbol{x})$，故据定义 1，可知$\mathscr{T}$是直交变换，即证得了条件的充分性.

反之，设$\mathscr{T}$为直交变换，则

$$\begin{aligned}(\boldsymbol{x}+\boldsymbol{y}, \boldsymbol{x}+\boldsymbol{y})&=((\boldsymbol{x}+\boldsymbol{y})\mathscr{T}, (\boldsymbol{x}+\boldsymbol{y})\mathscr{T})\\&=(\boldsymbol{x}\mathscr{T}+\boldsymbol{y}\mathscr{T}, \boldsymbol{x}\mathscr{T}+\boldsymbol{y}\mathscr{T});\end{aligned}$$

再展开左右两边，得

$$(x, x)+(y, y)+2\cdot(x, y)$$
$$=(x\mathscr{T}, x\mathscr{T})+(y\mathscr{T}, y\mathscr{T})+2\cdot(x\mathscr{T}, y\mathscr{T}),$$

因而，再利用 $(x, x)=(x\mathscr{T}, x\mathscr{T})$ 及 $(y, y)=(y\mathscr{T}, y\mathscr{T})$ 之关系，就可知 $(x, y)=(x\mathscr{T}, y\mathscr{T})$，这就证得了必要性. 定理 1 全部得证.

定理 2 设 $e_1, e_2, \cdots, e_n$ 是欧氏空间 $\mathfrak{R}$ 的标准直交基底. 线性变换 $\mathscr{T}$ 为直交变换的充要条件是 $e_1\mathscr{T}, e_2\mathscr{T}, \cdots, e_n\mathscr{T}$ 为标准直交基底.

证明 设 $\mathscr{T}$ 为直交变换，则由定理 1 得知 $(e_i\mathscr{T}, e_j\mathscr{T})=(e_i, e_j)=\delta_{ij}$，说明了 $e_1\mathscr{T}, e_2\mathscr{T}, \cdots, e_n\mathscr{T}$ 之标准直交性，即证得了必要性.

反之，如果 $e_1\mathscr{T}, e_2\mathscr{T}, \cdots, e_n\mathscr{T}$ 为标准直交基底，当 $x=\sum_{i=1}^{n}\xi_i e_i$，$y=\sum_{j=1}^{n}\eta_j e_j$ 时，由 $(e_i\mathscr{T}, e_j\mathscr{T})=\delta_{ij}$，有

$$(x\mathscr{T}, y\mathscr{T})=\left(\sum_{i=1}^{n}\xi_i\cdot(e_i\mathscr{T}), \sum_{j=1}^{n}\eta_j\cdot(e_j\mathscr{T})\right)$$
$$=\sum_{i,j=1}^{n}\xi_i\eta_j(e_i\mathscr{T}, e_j\mathscr{T})=\sum_{i=1}^{n}\xi_i\eta_i.$$

同样，由 $(e_i, e_j)=\delta_{ij}$ 又知 $(x, y)=\sum_{i=1}^{n}\xi_i\eta_i$. 因此，有 $(x\mathscr{T}, y\mathscr{T})=(x, y)$，即 $\mathscr{T}$ 是直交变换.

于是，定理 2 完全获证.

附注：定理 2 说明，欧氏空间 $\mathfrak{R}$ 中线性变换 $\mathscr{T}$ 为直交变换的充要条件是 $\mathscr{T}$ 使标准直交基底变为标准直交基底.

再来看 (8) 式中矩阵 Q 之特征. 由于 (6)、(7) 都是标准直交基底，故 $(x, x)=\sum\xi_i^2=\sum\eta_i^2$. 另方面，由 (9) 式则有

$$(\boldsymbol{x},\ \boldsymbol{x})=\sum\xi_i^2=(\xi_1,\ \cdots,\ \xi_n)\begin{pmatrix}\xi_1\\ \vdots\\ \xi_n\end{pmatrix}$$

$$=(\eta_1,\ \cdots,\ \eta_n)\cdot\boldsymbol{Q}\cdot\boldsymbol{Q}'\cdot\begin{pmatrix}\eta_1\\ \vdots\\ \eta_n\end{pmatrix},$$

故再从$(\boldsymbol{x},\ \boldsymbol{x})=\sum\eta_i^2=(\eta_1,\ \cdots,\ \eta_n)\cdot\begin{pmatrix}\eta_1\\ \vdots\\ \eta_n\end{pmatrix}$,就有

$$\boldsymbol{QQ}'=\boldsymbol{E}\quad \text{或}\quad \boldsymbol{Q}'=\boldsymbol{Q}^{-1}.$$

这说明了 $\boldsymbol{Q}$ 之转置矩阵等于 $\boldsymbol{Q}$ 之逆矩阵. 具有这样性质之矩阵 $\boldsymbol{Q}$,叫直交矩阵. 所以叫直交的原因,也是由于它是从标准直交基底所引起的. 即有:

定义 2 具有性质 $Q'=Q^{-1}$ 的(实)矩阵 $\boldsymbol{Q}$ 叫做直交矩阵.

不难证明直交矩阵的特征表现在下面的

定理 3 设 $\boldsymbol{A}=(a_{ij})$ 是一个 n 级(实)矩阵, 那末下面几个性质互为等价的:

i) $\boldsymbol{A}$ 是直交的(即 $\boldsymbol{A}'=\boldsymbol{A}^{-1}$),

ii) $\boldsymbol{AA}'=\boldsymbol{E}$,

iii) $\boldsymbol{A}'\boldsymbol{A}=\boldsymbol{E}$,

iv) $a_{i1}a_{j1}+a_{i2}a_{j2}+\cdots+a_{in}a_{jn}=\delta_{ij}$,

v) $a_{1i}a_{1j}+a_{2i}a_{2j}+\cdots+a_{ni}a_{nj}=\delta_{ij}$.

于是,又有:

推论 若 $\boldsymbol{A}$ 是直交矩阵, 则 $\boldsymbol{A}'$, $\boldsymbol{A}^*$, $\boldsymbol{A}^{-1}$ 也都是直交的; 又,若 $\boldsymbol{B}$ 是直交的,则 $\boldsymbol{AB}$ 也是直交的; 又,直交矩阵的行列式等于 ±1(行列式等于 1 的直交矩阵叫正常的, 等于 -1 的叫非正常的).

证明 从 $\boldsymbol{A}'=\boldsymbol{A}^{-1}$ 得 $\boldsymbol{A}=(\boldsymbol{A}^{-1})'$，即 $(\boldsymbol{A}^{-1})^{-1}=(\boldsymbol{A}^{-1})'$，即 $\boldsymbol{A}^{-1}$ 为直交的．再 $(\boldsymbol{A}')'=(\boldsymbol{A}^{-1})'=(\boldsymbol{A}')^{-1}$ 说明 $\boldsymbol{A}'$ 的直交性．又 $(\boldsymbol{A}^*)'=(\boldsymbol{A}')^*=(\boldsymbol{A}^{-1})^*=(\boldsymbol{A}^*)^{-1}$ 说明 $\boldsymbol{A}^*$ 之直交性．再有 $(\boldsymbol{AB})'=\boldsymbol{B}'\boldsymbol{A}'=\boldsymbol{B}^{-1}\boldsymbol{A}^{-1}=(\boldsymbol{AB})^{-1}$，则表明了 $\boldsymbol{AB}$ 的直交性．最后，从 $\boldsymbol{AA}'=\boldsymbol{E}$ 得 $1=|\boldsymbol{E}|=|\boldsymbol{AA}'|=|\boldsymbol{A}|\cdot|\boldsymbol{A}'|=|\boldsymbol{A}|^2$，故 $|\boldsymbol{A}|=\pm 1$．

直交变换与直交矩阵之间，有下面一个重要关系，即

定理 4 欧氏空间 $\mathfrak{R}$ 的线性变换 $\mathscr{T}$ 为直交变换的充要条件是它关于标准直交基底之矩阵为直交矩阵．

事实上，(8) 式已足够说明定理 4 之正确性．下面给出直接的证明．

证明 设 $\boldsymbol{e}_1, \boldsymbol{e}_2, \cdots, \boldsymbol{e}_n$ 为欧氏空间 $\mathfrak{R}$ 之标准直交基底．线性变换 $\mathscr{T}$ 关于这组基底之矩阵是 $\boldsymbol{A}=(a_{ij})$，即

$$\begin{pmatrix} \boldsymbol{e}_1\mathscr{T} \\ \boldsymbol{e}_2\mathscr{T} \\ \vdots \\ \boldsymbol{e}_n\mathscr{T} \end{pmatrix}=\boldsymbol{A}\cdot\begin{pmatrix} \boldsymbol{e}_1 \\ \boldsymbol{e}_2 \\ \vdots \\ \boldsymbol{e}_n \end{pmatrix},$$

或

$$\boldsymbol{e}_i\mathscr{T}=a_{i1}\boldsymbol{e}_1+a_{i2}\boldsymbol{e}_2+\cdots+a_{in}\boldsymbol{e}_n,\quad (i=1, 2, \cdots, n).$$

于是，

$$\begin{aligned}(\boldsymbol{e}_i\mathscr{T}, \boldsymbol{e}_j\mathscr{T})&=\left(\sum_{s=1}^{n} a_{is}\boldsymbol{e}_s,\ \sum_{t=1}^{n} a_{jt}\boldsymbol{e}_t\right)\\ &=\sum_{s,t=1}^{n} a_{is}a_{jt}\cdot(\boldsymbol{e}_s, \boldsymbol{e}_t)\\ &=\sum_{s,t=1}^{n} a_{is}a_{jt}\delta_{st}=\sum_{t=1}^{n} a_{it}a_{jt}.\end{aligned}\tag{10}$$

若 $\mathscr{T}$ 为直交变换，则由定理 2 得知 $\boldsymbol{e}_1\mathscr{T}, \cdots, \boldsymbol{e}_n\mathscr{T}$ 为 $\mathfrak{R}$ 之标准直交基底，因之 $(\boldsymbol{e}_i\mathscr{T}, \boldsymbol{e}_j\mathscr{T})=\delta_{ij}$，故从 (10) 式得

$$\sum_{t=1}^{n} a_{it}a_{jt}=\delta_{ij},$$

于是，据定理 3，即知 $\boldsymbol{A}$ 是直交矩阵.

反之，设 $\boldsymbol{A}$ 为直交矩阵，于是，据定理 3 的 iv)，得知 $\sum_{t=1}^{n} a_{it}a_{jt}=\delta_{ij}$，故从 (10) 得 $(\boldsymbol{e}_i\mathscr{T}, \boldsymbol{e}_j\mathscr{T})=\delta_{ij}$，即 $\boldsymbol{e}_1\mathscr{T}, \boldsymbol{e}_2\mathscr{T}, \cdots, \boldsymbol{e}_n\mathscr{T}$ 为 $\mathfrak{R}$ 之标准直交基底，因而由定理 2 得知 $\mathscr{T}$ 为直交变换.

定理 4 完全获证.

定理 4 表明：直交变换与直交矩阵之间有着一一对应的关系. 但是，这种一一对应关系是建立在标准直交基底上的，如果基底不是标准直交基底，那末定理 4 就不再成立了，这由下面的两个例子可看出.

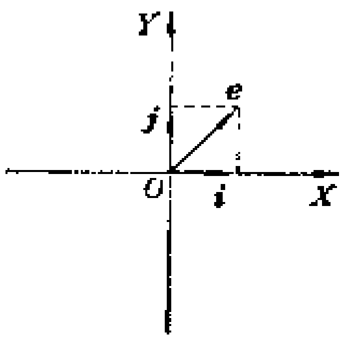


图 7-3

[例 1] 考察 $\mathfrak{R}_2$ 中以 x-轴为对称轴的反射变换 $\mathscr{T}$.

显然，反射变换 $\mathscr{T}$ 是使 $\mathfrak{R}_2$ 中每个向量的长都不变，所以 $\mathscr{T}$ 是 $\mathfrak{R}_2$ 的直交变换. 假如以基本单位向量 $\boldsymbol{i}, \boldsymbol{j}$ 为 $\mathfrak{R}_2$ 之基底，如图 7-3，则易知 $\mathscr{T}$ 对应的矩阵 $\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}$ 是直交的（因为 $\boldsymbol{i}\mathscr{T}=\boldsymbol{i}$，$\boldsymbol{j}\mathscr{T}=-\boldsymbol{j}$）；但如果采用 $\boldsymbol{e}$ 与 $\boldsymbol{j}$ 为基底，$\boldsymbol{e}$ 是以 $\boldsymbol{i}, \boldsymbol{j}$ 为边的正方形的对角线向量，则 $\boldsymbol{e}=\boldsymbol{i}+\boldsymbol{j}$，而 $\boldsymbol{e}$ 与 $\boldsymbol{j}$ 不再是标准直交基底，这时因 $\boldsymbol{e}\mathscr{T}=\boldsymbol{i}-\boldsymbol{j}$，故

$$\begin{pmatrix}\boldsymbol{e}\mathscr{T}\\ \boldsymbol{j}\mathscr{T}\end{pmatrix}=\begin{pmatrix}\boldsymbol{i}-\boldsymbol{j}\\ -\boldsymbol{j}\end{pmatrix}=\begin{pmatrix}\boldsymbol{e}-2\boldsymbol{j}\\ -\boldsymbol{j}\end{pmatrix}=\begin{pmatrix}1 & -2\\ 0 & -1\end{pmatrix}\begin{pmatrix}\boldsymbol{e}\\ \boldsymbol{j}\end{pmatrix},$$

即反射变换 $\mathscr{T}$（直交的）对应的矩阵 $\begin{pmatrix}1 & -2\\ 0 & -1\end{pmatrix}$ 不再是直交

矩阵.

［例 2］ 还是考察 $\mathfrak{R}_2$, 如图 7-4, 令向量 $\boldsymbol{e}$ 之意义同例 1 中的一样. 今作 $\mathfrak{R}_2$ 的线性变换 $\mathscr{S}$, 它具有性质 $\boldsymbol{e}\mathscr{S}=-\boldsymbol{e}$, $\boldsymbol{j}\mathscr{S}=\boldsymbol{j}$. 因之, 由

$$\begin{pmatrix}\boldsymbol{e}\mathscr{S}\\ \boldsymbol{j}\mathscr{S}\end{pmatrix}=\begin{pmatrix}-\boldsymbol{e}\\ \boldsymbol{j}\end{pmatrix}=\begin{pmatrix}-1 & 0\\ 0 & 1\end{pmatrix}\begin{pmatrix}\boldsymbol{e}\\ \boldsymbol{j}\end{pmatrix},$$

则知 $\mathscr{S}$ 关于基底 $\boldsymbol{e}, \boldsymbol{j}$ 之矩阵 $\begin{pmatrix}-1 & 0\\ 0 & 1\end{pmatrix}$ 为直交的; 但变换 $\mathscr{S}$ 确非直交: 这是因为, 从初等几何学知道:

$$(\boldsymbol{j}+\boldsymbol{e})\mathscr{S}=\boldsymbol{j}\mathscr{S}+\boldsymbol{e}\mathscr{S}=\boldsymbol{j}-\boldsymbol{e}=-\boldsymbol{i},\ \boldsymbol{j}+\boldsymbol{e}=\boldsymbol{f},$$

$$\|\boldsymbol{j}+\boldsymbol{e}\|^2=\|\boldsymbol{f}\|^2=\|\boldsymbol{e}\|^2+\|\boldsymbol{j}\|^2-2\cdot\|\boldsymbol{e}\|\cdot\|\boldsymbol{j}\|\cdot\cos\frac{3\pi}{4}$$

$$=2+1-2\sqrt{2}\left(-\frac{1}{\sqrt{2}}\right)=5,$$

$$\|(\boldsymbol{j}+\boldsymbol{e})\mathscr{S}\|=\|-\boldsymbol{i}\|=1,$$

所以 $\|\boldsymbol{f}\mathscr{S}\|\neq\|\boldsymbol{f}\|$, 因此 $\mathscr{S}$ 不是直交变换.

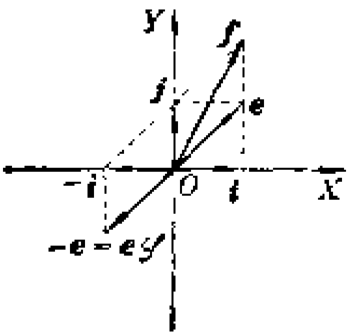


图 7-4

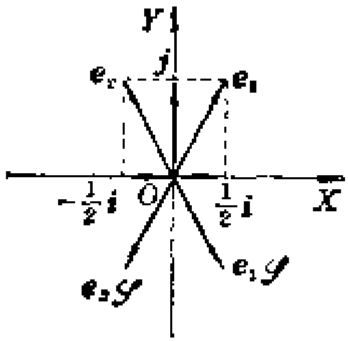


图 7-5

例 1 与例 2 说明: 当基底不是标准直交时, 直交变换对应的矩阵可以不是直交矩阵, 而直交矩阵对应的线性变换也可以不是直交变换. 但是, 虽然基底不是标准直交, 而直交变换所对应的矩阵也有可能是直交的, 下面例 3 就说明这问题

［例 3］ 同例 1 一样, 仍考察平面 $\mathfrak{R}_2$ 中以 x-轴为对称

轴的反射变换 $\mathscr{S}$. 如图 7-5, 令 $e_1=\frac{1}{2}i+j$, $e_2=-\frac{1}{2}i+j$, 于是 e_1 与 e_2 对于 y 轴是对称的. 因 e_1, e_2 不直交, 故 e_1, e_2 不是 $\mathfrak{R}_2$ 的标准直交基底, 但 $e_1\mathscr{S}=\frac{1}{2}i-j$, $e_2\mathscr{S}=-\frac{1}{2}i-j$, 所以 $e_1\mathscr{S}=-e_2$, $e_2\mathscr{S}=-e_1$, 于是 $\begin{pmatrix}e_1\mathscr{S}\\e_2\mathscr{S}\end{pmatrix}=\begin{pmatrix}0&-1\\-1&0\end{pmatrix}\cdot\begin{pmatrix}e_1\\e_2\end{pmatrix}$, 这说明了虽然基底 e_1、e_2 不是标准直交基底, 但直交变换 $\mathscr{S}$ 所对应的矩阵 $\begin{pmatrix}0&-1\\-1&0\end{pmatrix}$ 还是直交的.

练习三十二

1. 证明: 欧氏空间 $\mathfrak{R}$ 中的直交变换 $\mathscr{T}$, 使任二个向量间的夹角都保持不变, 即 $(\widehat{x, y})=(\widehat{x\mathscr{T}, y\mathscr{T}})$. 并以相似变换为例, 说明它的逆定理不成立.

2. $\mathscr{T}$ 为直交变换的充要条件是使任二点间的距离保持不变, 即 $\|x-y\|=\|x\mathscr{T}-y\mathscr{T}\|$.

3. 如果 A 是直交矩阵, 证明当 P 是直交矩阵时, $P^{-1}AP$ 也是直交的. 试举一例说明 A 为直交矩阵时, 虽 P 不是直交, 但 $P^{-1}AP$ 也可能还是直交的.

4. 设 A 为直交矩阵. 证明: $A^{-1}SA$ 是对称或反对称, 全由 S 为对称或反对称而定.

5. 证明: 直交矩阵的任一行(列)乘以(-1), 或任二行(列)互换, 结果仍是直交矩阵.

6. 证明下列三性质中任一个是另二个的结果: i) A 是对称的, ii) A 是直交的, iii) $A^2=E$.

7. 直交矩阵中每元素或等于它的代数余子式, 或等于其代数余子式之反号, 全由这矩阵为正常或非正常来决定.

8. n 维欧氏空间中至多有 $n+1$ 个向量, 使两两间的夹角都大于 $90°$.

第4节 主轴问题

二维空间 $\mathfrak{R}_2$ 中直角坐标轴旋转变换的矩阵

$\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$为直交矩阵，三维空间 $\mathfrak{R}_3$ 中直角坐标轴旋转变换的矩阵$\begin{pmatrix} \lambda_1 & \lambda_2 & \lambda_3 \\ \mu_1 & \mu_2 & \mu_3 \\ \nu_1 & \nu_2 & \nu_3 \end{pmatrix}$也是直交矩阵，而 n 维欧氏空间 $\mathfrak{R}$ 中由一组标准直交基底变为另一组标准直交基底的变换矩阵也是直交矩阵．正因为这样，欧氏空间中由标准直交基底变为另一组标准直交基底的直交矩阵可以认为是二维、三维空间中直角坐标轴旋转变换矩阵的直接推广．

我们又知道，二维空间 $\mathfrak{R}_2$ 中，中心在坐标原点的二次有心曲线方程 $Ax^2+2Bxy+Cy^2=F$ 之左端，可以借坐标轴的旋转变换

$$\begin{pmatrix} x \\ y \end{pmatrix}=\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}\begin{pmatrix} x' \\ y' \end{pmatrix}$$

化为平方和 $A'x'^2+C'y'^2$ 的形状．关于三维空间中有心二次曲面，也有类似的结论（我们不打算象解析几何学里那样用烦杂的运算来解决这个问题，留在后面讲二次曲面的度量分类时用直交变换的观点一并来解决）．于是，推广到 n 个变量上，就是二次齐式 $\sum a_{ij}x_ix_j=\boldsymbol{X}'\boldsymbol{A}\boldsymbol{X}$ 也可以借助直交矩阵 $\boldsymbol{Q}$ 的变换

$$\boldsymbol{X}=\boldsymbol{Q}\boldsymbol{Y}$$

使之变成平方和，这就是我们早在第六章里提出的所谓主轴问题．虽然第六章里已把二次齐式 $\sum a_{ij}x_ix_j$ 化为平方和，即要求

$$\boldsymbol{Q}'\boldsymbol{A}\boldsymbol{Q}$$

为对角矩阵，但那里的 $\boldsymbol{Q}$ 没有附加任何限制，可是现在要求

$Q'=Q^{-1}$. 如果存在这样的 Q, 使

$$Q'AQ=Q^{-1}AQ=\mathrm{diag}\{d_1, d_2, \cdots, d_n\},$$

则由第五章可知 $d_1, d_2, \cdots, d_n$ 必是 A 的 n 个特征值. 因而, 首先需要有下面的:

引理 1 实对称矩阵之特征值为实数.

证明 设 $A=(a_{ij})$ 是实对称矩阵. 若 α 为 A 之一个特征值, 命 X 是 A 对应于 α 的一个特征向量, 则 $AX=\alpha X$, 于是

$$\bar{\alpha}\overline{X}'=\overline{X}'\overline{A}'=\overline{X}'A'=\overline{X}'A.$$

因为 $\overline{A}=A$, $A'=A$, 所以

$$\bar{\alpha}\overline{X}'X=\overline{X}'AX=\alpha\overline{X}'X,$$

因此 $(\alpha-\bar{\alpha})\cdot\overline{X}'X=0$, 但 $X\neq 0$, 所以 $\overline{X}'X=\sum x_i\bar{x}_i>0$, 故 $\alpha=\bar{\alpha}$, 即 α 是实数, 证完.

注意: 引理 1 中的对称矩阵如果不是实矩阵, 结论就不成立. 例如 $A=\begin{pmatrix} i & 1 \\ 1 & 0 \end{pmatrix}$ 为对称矩阵, 但其特征方程 $|\lambda E-A|=\begin{vmatrix} \lambda-i & -1 \\ -1 & \lambda \end{vmatrix}=\lambda^2-i\lambda-1=0$ 显然没有实根.

有了引理 1, 就能够解决我们的主要问题, 即:

定理 1 设 A 是实对称矩阵 ($\overline{A}=A$, $A'=A$), 则至少有一个实直交矩阵 P ($\overline{P}=P$, $P'=P^{-1}$), 使 $P'AP=P^{-1}AP$ 为对角矩阵.

证明 令 λ_1 为 A 的一个特征值, 则据引理 1 知 λ_1 是实数, 故由 $AX=\lambda_1 X$ 是具实系数的线性方程组, 可知对应于 λ_1 有一个实特征向量 $X_1\neq 0$ 使 $AX_1=\lambda_1 X_1$, 且可令 $\|X_1\|=$

$\sqrt{\sum x_i^2}=1$†(这是因为，若$\|X_1\|\neq 1$，其标准化 $Y_1=\frac{1}{\|X_1\|}X_1$ 即为对应于 λ_1 之特征向量)．

假如能够作一个(实)直交矩阵 P_1，使 X_1 为它的第一个列向量，而令 P_1 之其他列向量分别用 $X_2, \cdots, X_n$ 表示，则由直接利用矩阵乘法规则，容易看出 $P_1^{-1}AP_1=(P_1^{-1}A)P_1$ 之第一个列向量是

$$(P_1^{-1}A)X_1=P_1^{-1}(AX_1)=\lambda_1 P_1^{-1}X_1=\lambda_1\cdot\begin{pmatrix}1\\0\\\vdots\\0\end{pmatrix},$$

这是因为从 $P_1^{-1}P_1=E$ 即得 $P_1^{-1}X_1=\begin{pmatrix}1\\0\\\vdots\\0\end{pmatrix}$．因而，有

$$P_1^{-1}AP_1=\begin{pmatrix}\lambda_1 & s_2 & \cdots & s_n\\0 & & & \\\vdots & & B & \\0 & & & \end{pmatrix},$$

这里 B 是 $n-1$ 级实矩阵．

但$(P_1^{-1}AP_1)'=P_1'A'(P_1^{-1})'=P_1^{-1}AP_1$，这说明$P_1^{-1}AP_1$为对称的，故 $s_2=\cdots=s_n=0$，$B'=B$．

今用归纳法来证明．$n=1$ 时定理显然正确，因为这时 P_1 与 A 是实数且 $P_1=1$．假定对 $n-1$ 级矩阵的情况定理成立，即有 $n-1$ 级的一个(实)直交矩阵 Q 使

† 由 n 实数组而成之空间 $\mathfrak{R}_n$，恒定义 $\alpha=(\alpha_1, \cdots, \alpha_n)$ 与 $\beta=(\beta_1,\cdots,\beta_n)$ 之数积 $(\alpha, \beta)=\sum\alpha_i\beta_i$ 使 $\mathfrak{R}_n$ 为欧氏空间，因而 $\|\alpha\|=\sqrt{\sum\alpha_i^2}$，都是对标准直交基底 $e_1=(1, 0, \cdots, 0), \cdots, e_n=(0, \cdots, 0, 1)$来说的．

$$Q^{-1}BQ=\begin{pmatrix}\lambda_2 & & \\ & \ddots & \\ & & \lambda_n\end{pmatrix}=\text{diag}\{\lambda_2, \cdots, \lambda_n\}.$$

于是，令
$$P_2=\begin{pmatrix}1 & 0 & \cdots & 0\\ 0 & & & \\ \vdots & & Q & \\ 0 & & & \end{pmatrix},$$

则
$$P_2^{-1}=\begin{pmatrix}1 & 0 & \cdots & 0\\ 0 & & & \\ \vdots & & Q^{-1} & \\ 0 & & & \end{pmatrix}=\begin{pmatrix}1 & \\ & Q'\end{pmatrix}=\begin{pmatrix}1 & \\ & Q\end{pmatrix}'=P_2',$$

即 P_2 为 n 级(实)直交矩阵，因之 $P=P_1P_2$ 是(实)直交的．由是，

$$P^{-1}AP=P_2^{-1}(P_1^{-1}AP_1)P_2=\begin{pmatrix}1 & \\ & Q^{-1}\end{pmatrix}\begin{pmatrix}\lambda_1 & \\ & B\end{pmatrix}\begin{pmatrix}1 & \\ & Q\end{pmatrix}$$
$$=\begin{pmatrix}\lambda_1 & \\ & Q^{-1}BQ\end{pmatrix}=\text{diag}\{\lambda_1, \lambda_2, \cdots, \lambda_n\}.$$

因此如果能证明下面的引理 2，则定理 1 就完全解决了．

引理 2 设有由 n 实数组而成之向量 $X_1=\begin{pmatrix}\xi_1\\ \vdots\\ \xi_n\end{pmatrix}$，它的长 $\|X_1\|=\sqrt{(X_1, X_1)}=\sqrt{\sum\xi_i^2}=1$，那就至少有一个实直交矩阵 P，它的第一个列向量为 X_1，且行列式 $|P|$ 可任取 $+1$ 或 -1．

证明 因 $X_1\neq 0$，故一定可选取 $n-1$ 个 n 实数组的向量

$$Y_i=\begin{pmatrix}a_{1i}\\ a_{2i}\\ \vdots\\ a_{ni}\end{pmatrix}, \quad (i=2, \cdots, n)$$

使得行列式 $\begin{vmatrix} \xi_1 & a_{12} & \cdots & a_{1n} \\ \xi_2 & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ \xi_n & a_{n2} & \cdots & a_{nn} \end{vmatrix} \neq 0.$

这样的一组 $\boldsymbol{Y}_i$ 不仅存在，而且选取的方法很多．于是 $\boldsymbol{X}_1, \boldsymbol{Y}_2, \cdots, \boldsymbol{Y}_n$ 为由所有 n 个实数组形成的欧氏空间 $\Re_n$ 之一组基底；再利用第 2 节定理 2 之证明方法，将 $\boldsymbol{X}_1, \boldsymbol{Y}_2, \cdots, \boldsymbol{Y}_n$ 标准直交化，就得到 n 个两两直交的实单位向量 $\boldsymbol{X}_1, \boldsymbol{X}_2, \cdots, \boldsymbol{X}_n$．命 $\boldsymbol{P}$ 是以 $\boldsymbol{X}_1, \boldsymbol{X}_2, \cdots, \boldsymbol{X}_n$ 为 n 个列向量之矩阵，由于 $(\boldsymbol{X}_i, \boldsymbol{X}_j)=\delta_{ij}$，则据前节定理 3 之 v)，得知 $\boldsymbol{P}$ 是以 $\boldsymbol{X}_1$ 为第一个列向量的直交矩阵．再因 $|\boldsymbol{P}|=\pm 1$，且调换 $\boldsymbol{P}$ 中二列不影响 $\boldsymbol{P}$ 之直交性而只改变 $|\boldsymbol{P}|$ 之正负号，故我们所需要的 $\boldsymbol{P}$ 之行列式 $|\boldsymbol{P}|$ 可为 $+1$ 也可为 -1．引理 2 证完．

引理 2 中的直交矩阵 $\boldsymbol{P}$，可按下法求出：首先，考虑 n 个未知量 $x_1, \cdots, x_n$ 的齐次方程

$$\boldsymbol{X}'_1\boldsymbol{X}=\xi_1x_1+\xi_2x_2+\cdots+\xi_nx_n=0.$$

取其任一个非零实解 $\boldsymbol{Y}_2$，则其标准化 $\boldsymbol{X}_2=\boldsymbol{Y}_2/\sqrt{\boldsymbol{Y}'_2\boldsymbol{Y}_2}$ 为一个实单位向量，且和 $\boldsymbol{X}_1$ 直交．

再考虑两个线性无关的齐次方程（n 个未知量 $x_1, \cdots, x_n$ 的）：

$$\boldsymbol{X}'_1\cdot\boldsymbol{X}=0, \quad \boldsymbol{X}'_2\cdot\boldsymbol{X}=0,$$

则其任一个非零（实）解 $\boldsymbol{Y}_3$ 之标准化

$$\boldsymbol{X}_3=\boldsymbol{Y}_3/\sqrt{\boldsymbol{Y}'_3\boldsymbol{Y}_3}$$

为一个实单位向量，且和 $\boldsymbol{X}_1$ 与 $\boldsymbol{X}_2$ 都是直交的．

继续这样作下去，假定已求得 $n-1$ 个两两直交的实单位向量 $\boldsymbol{X}_1, \cdots, \boldsymbol{X}_{n-1}$；如果 $\boldsymbol{Y}_n$ 是齐次方程组

$$\boldsymbol{X}'_1\boldsymbol{X}=0, \ \boldsymbol{X}'_2\boldsymbol{X}=0, \ \cdots, \ \boldsymbol{X}'_{n-1}\boldsymbol{X}=0$$

之任一个非零(实)解，则其标准化

$$X_n = Y_n / \sqrt{Y_n' \cdot Y_n}$$

是和 $X_1, \cdots, X_{n-1}$ 都直交的实单位向量. 这样，矩阵 $P = (X_1, \cdots, X_n)$ 即为所求之实直交矩阵.

定理 1 实质上就是主轴问题，不过它是用矩阵的形式表达的，如用几何的语言来叙述，就是：

定理 2（主轴问题）. 设在 n 维欧氏空间 $\mathfrak{R}$ 中，取定一组标准直交基底 $e_1, e_2, \cdots, e_n$. 如果

$$\varphi(\xi_1, \xi_2, \cdots, \xi_n) = \sum a_{ij}\xi_i\xi_j = X'AX$$

$$[\text{这里 } X' = (\xi_1, \cdots, \xi_n)]$$

是欧氏空间 $\mathfrak{R}$ 在基底 $e_1, \cdots, e_n$ 内之二次齐式，而 $\xi_1, \cdots, \xi_n$ 表示 $\mathfrak{R}$ 之向量 x 关于 $e_1, \cdots, e_n$ 的坐标，即 $x = \xi_1 e_1 + \xi_2 e_2 + \cdots + \xi_n e_n$，那末必有直交变换 $\mathscr{T}$ 把 $e_1, e_2, \cdots, e_n$ 变到另一组标准直交基底 $f_1, f_2, \cdots, f_n$，使在这新基底内所给的二次齐式 $\sum a_{ij}\xi_i\xi_j = X'AX$ 变为典型式 $\sum \lambda_i \eta_i^2 = Y' \cdot \mathrm{diag}\{\lambda_1, \lambda_2, \cdots, \lambda_n\} \cdot Y$，这里 $\eta_1, \cdots, \eta_n$ 是向量 x 关于基底 $f_1, \cdots, f_n$ 之坐标：$x = \eta_1 f_1 + \eta_2 f_2 + \cdots + \eta_n f_n$，$Y' = (\eta_1, \eta_2, \cdots, \eta_n)$.

证明 由定理 1 知，有直交矩阵 P，使

$$P'AP = P^{-1}AP = \mathrm{diag}\{\lambda_1, \lambda_2, \cdots, \lambda_n\}.$$

再令 $\begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix} \mathscr{T} = P' \begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix}$，据上节定理 4 与定理 2，得知 $\mathscr{T}$ 为直交变换，而 $f_1 = e_1\mathscr{T}, f_2 = e_2\mathscr{T}, \cdots, f_n = e_n\mathscr{T}$ 为 $\mathfrak{R}$ 之另一组标准直交基底. 容易验证 $X = PY$，故

$$X'AX = Y'P'APY = Y' \cdot \mathrm{diag}\{\lambda_1, \lambda_2, \cdots, \lambda_n\} \cdot Y,$$

即定理成立.

主轴问题解决了 n 个变量 $\xi_1, \xi_2, \cdots, \xi_n$ 的二次齐式

$\sum a_{ij}\xi_i\xi_j$ 总是可以借直交矩阵 $\boldsymbol{Q}$ 的变量间的线性变换

$$\begin{pmatrix}\xi_1\\ \vdots\\ \xi_n\end{pmatrix}=\boldsymbol{Q}\cdot\begin{pmatrix}\eta_1\\ \vdots\\ \eta_n\end{pmatrix}$$

使 $\sum a_{ij}\xi_i\xi_j=\sum\lambda_i\eta_i^2$. 故特别当 $n=2$ 时，即是说平面上恒有坐标轴的适当的旋转变换，使

$$Ax^2+2Bxy+Cy^2=A'x'^2+C'y'^2. \tag{11}$$

但是，主轴问题起源于解析几何学中有心二次曲线的理论，而有心二次曲线常局限于椭圆和双曲线两种；在解析几何中，

$$Ax^2+2Bxy+Cy^2=F \tag{12}$$

表示椭圆型与双曲线型的充要条件分别是 $B^2-AC<0$ 与 $B^2-AC>0$. 这就是说，问题是在 $B^2-AC\neq0$ 的情况下产生的，可是等到主轴问题解决了之后，再回到 $n=2$ 时，不论 B^2-AC 为零与否，则知恒有 (11) 这个关系式．也就是说，虽(12)中的 $B^2-AC=0$(这条件表示(12)为抛物线型的)，也会有(11)式． 这也是说，在 $B^2-AC=0$ 时，对新坐标轴，曲线(12)化为

$$A'x'^2+C'y'^2=F \tag{13}$$

的形式．然而(12)在 $B^2-AC=0$ 时表示抛物线的退化状况，即为二个平行的直线（实或虚或重合），这是由于在 $B^2-AC=0$ 时 $Ax^2+2Bxy+Cy^2$ 是一个完全平方式如 $(\alpha x+\beta y)^2$ 的缘故．可是，经过坐标轴的旋转变换后曲线的方程变成(13)，乍看起来好象是椭圆或双曲线，发生了矛盾．其实不然，因为我们知道定理 1 中对角矩阵 $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}$ 之对角线上的数都是矩阵 $\boldsymbol{A}$ 的特征值（实数），故(11)中之 A', C' 应该是对称矩阵 $\begin{pmatrix}A & B\\ B & C\end{pmatrix}$ 的特征值，但 $\begin{pmatrix}A & B\\ B & C\end{pmatrix}$ 在 $B^2-AC=0$ 时是降秩矩阵，

所以它必有一个特征值等于零，例如说 $C'=0$，于是(11)式实际上为

$$Ax^2+2Bxy+Cy^2=A'x'^2,$$

所以曲线方程 $Ax^2+2Bxy+Cy^2=F$ 简化为 $A'x'^2=F$，亦即这时曲线仍然表示平行的两个直线(实或虚，重合或相异)，和未简化方程时二次曲线表示的几何图形是一致的. 这是应提请读者注意的.

因为定理 2 中典型式 $\sum\lambda_i\eta_i^2$ 之系数 $\lambda_1, \cdots, \lambda_n$ 是定理 1 中对角矩阵 $\boldsymbol{P}'\boldsymbol{A}\boldsymbol{P}=\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}$ 的对角线上的数，因之也就是对称矩阵 $\boldsymbol{A}$ 的 n 个特征值(实数)，所以由第六章惯性定理及有定与不定齐式之理论，又可得到下面的两个定理.

定理 3 二次齐式 $\sum a_{ij}\xi_i\xi_j$ 的秩等于对称矩阵 $\boldsymbol{A}=(a_{ij})$ 中不为零之特征值的个数，它的正(负)惯性指数等于正(负)特征值之个数，它的正负号差等于正特征值之个数减去负特征值之个数.

定理 4 二次齐式 $\sum a_{ij}\xi_i\xi_j$ 为正(负)定的充要条件是它的 n 个特征值都是正(负)数；为半正定(半负定)的充要条件是 n 个特征值中有正数和零(有负数和零)；为不定的充要条件是 n 个特征值中正负数都有.

注意：定理 1 实际上还说明了实对称矩阵与对角矩阵相似. 因而，实对称矩阵的最小多项式无重根，或它的特征矩阵之初等因子都是一次式. 总之，在第五章里所讨论与对角矩阵相似之矩阵的有关结果，对于实对称矩阵都成立.

最后谈一下化二次齐式到主轴上去的直交矩阵的实际求法.

在第五章里曾经讲过：只要矩阵 $\boldsymbol{A}$ 能与对角矩阵相似，如 $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}=\mathrm{diag}\{\lambda_1, \cdots, \lambda_n\}$，不仅 $\lambda_i(i=1, \cdots, n)$ 是 $\boldsymbol{A}$ 之

特征值，而且 $\boldsymbol{P}$ 的第 i 个列向量还是对应于 λ_i 的特征向量；又知道这时任一个特征值的重数等于它的特征子空间的维数，也就是说这时对应于 $\boldsymbol{A}$ 之 k 重特征值的特征向量间线性无关的最大个数等于 k，故当 $\boldsymbol{A}$ 为实对称矩阵时，因其特征值全为实数，所以特征向量也都可取为实向量，于是定理 1 中实直交矩阵 $\boldsymbol{P}$ 表面上看可以这样来求得：即先求出 $\boldsymbol{A}$ 的每个特征值所对应的最大无关之特征向量组，因而得到了 $\boldsymbol{A}$ 的 n 个线性无关的特征（实）向量，再按照第 2 节中所讲的把它们标准直交化，就得到所要求的实直交矩阵 $\boldsymbol{P}$ 之 n 个列向量了．可是，将 n 个无关的向量标准直交化，当 n 相当大时已经是一个很费事的工作，问题的关键还不仅仅在于费事这一点，主要还在于标准直交化了之后，得到的实直交矩阵 $\boldsymbol{P}$ 之 n 个列向量到底还能不能为 $\boldsymbol{A}$ 之特征向量？于是求化二次齐式到主轴上去的实直交矩阵，到底用怎样的实际方法去求呢？为此，首先要证明下面的

引理 3 对应于实对称矩阵的两个不同特征值之特征向量（可取为实向量）是直交的．

在第五章里已证过：任何矩阵的两个不同特征值所对应的特征向量是线性无关的，但不一定直交；可是，对于实对称矩阵，它们不仅线性无关，而且还是直交的．

证明 设 $\boldsymbol{A}$ 实对称（即 $\overline{\boldsymbol{A}}=\boldsymbol{A}$, $\boldsymbol{A}'=\boldsymbol{A}$），$\lambda_1$ 与 λ_2 是 $\boldsymbol{A}$ 之两个不同（即 $\lambda_1\neq\lambda_2$）的特征值——因而都是实数，令 $\boldsymbol{X}_1$ 与 $\boldsymbol{X}_2$ 是分别对应于 λ_1 与 λ_2 的（实）特征向量，即

$$\boldsymbol{A}\boldsymbol{X}_1=\lambda_1\boldsymbol{X}_1,\qquad \boldsymbol{A}\boldsymbol{X}_2=\lambda_2\boldsymbol{X}_2,$$

如果 $\boldsymbol{X}_1=\begin{pmatrix}\xi_1\\ \vdots\\ \xi_n\end{pmatrix}$，$\boldsymbol{X}_2=\begin{pmatrix}\eta_1\\ \vdots\\ \eta_n\end{pmatrix}$，我们的问题就是要证明

$$X_1'X_2=(\xi_1,\ \cdots,\ \xi_n)\cdot\begin{pmatrix}\eta_1\\ \vdots\\ \eta_n\end{pmatrix}=\sum\xi_i\eta_i=0.$$

因为 $X_1'A'=\lambda_1X_1'$，故 $X_1'\cdot A'X_2=\lambda_1X_1'X_2$，而 $A'=A$，从而 $X_1'AX_2=\lambda_1X_1'X_2$．于是，再以 λ_2X_2 去替换 AX_2，就得到 $X_1'\cdot\lambda_2X_2=\lambda_1X_1'X_2$，即 $(\lambda_1-\lambda_2)\cdot X_1'X_2=0$．但 $\lambda_1\neq\lambda_2$，故 $X_1'X_2=0$，证完．

于是把上面所说的归纳起来，则知主轴问题的实际解法可以归结为下面几个具体步骤：

1) 首先求二次齐式 $\sum a_{ij}\xi_i\xi_j$ 的实对称矩阵 $A=(a_{ij})$ 之 n 个特征值 $\lambda_1,\ \lambda_2,\ \cdots,\ \lambda_n$（都是实的）．于是，$A\sim\mathrm{diag}\{\lambda_1,\ \lambda_2,\ \cdots,\ \lambda_n\}$，因此 $\sum a_{ij}\xi_i\xi_j$ 之典型式为

$$\lambda_1\eta_1^2+\lambda_2\eta_2^2+\cdots+\lambda_n\eta_n^2,$$

即 $\sum a_{ij}\xi_i\xi_j=\lambda_1\eta_1^2+\lambda_2\eta_2^2+\cdots+\lambda_n\eta_n^2$．

如果不仅是要求典型式，还需求出化为典型式的（实）直交变换，也就是说要求出使

$$P^{-1}AP=\mathrm{diag}\{\lambda_1,\ \lambda_2,\ \cdots,\ \lambda_n\}$$

之实直交矩阵 P，那就还需进行下面的手续．

2) i) 设 $\lambda_1,\ \cdots,\ \lambda_t$ 是 A 之不同的特征值，其重数分别令为 $k_1,\ k_2,\ \cdots,\ k_t$，于是 $k_1+k_2+\cdots+k_t=n$．对于每个 $\lambda_i(i=1,\ 2,\ \cdots,\ t)$，求出线性齐次方程组 $(\lambda_iE-A)X=0$ 的基础解系 $\gamma_{i1},\ \gamma_{i2},\ \cdots,\ \gamma_{ik_i}$，即为 k_i 个线性无关的（实）特征向量．这样，我们共有 n 个线性无关的（实）特征向量：

$$\gamma_{11},\ \cdots,\ \gamma_{1k_1};\ \gamma_{21},\ \cdots,\ \gamma_{2k_2};\ \cdots\cdots;\ \gamma_{t1},\ \cdots,\ \gamma_{tk_t}.$$

ii) 把每一组特征向量 $\gamma_{i1},\ \cdots,\ \gamma_{ik_i}$ 象第 2 节中所说的那样加以标准直交化，得到 k_i 个两两直交的单位向量 X_{i1}，

$\cdots$, X_{ik_i}, 这样共有 n 个两两直交的单位向量:

$$X_{11},\ \cdots,\ X_{1k_1};\ X_{21},\ \cdots,\ X_{2k_2};\ \cdots\cdots;\ X_{t1},\ \cdots,\ X_{tk_t}.$$

iii) 以 X_{11}, $\cdots$, X_{1k_1}, $\cdots\cdots$, X_{t1}, $\cdots$, X_{tk_t} 顺次作为列向量作矩阵 P, 这 P 就是我们所要求的实直交矩阵.

[例] 化二次齐式 $\varphi(x_1, x_2, x_3, x_4)=2x_1x_2+2x_1x_3-2x_1x_4-2x_2x_3+2x_2x_4+2x_3x_4$ 到主轴上去.

这时,齐式的实对称矩阵是

$$A=\begin{pmatrix} 0 & 1 & 1 & -1 \\ 1 & 0 & -1 & 1 \\ 1 & -1 & 0 & 1 \\ -1 & 1 & 1 & 0 \end{pmatrix}.$$

于是 $|\lambda E-A|=(\lambda-1)^3(\lambda+3)$, 故 A 有 1 为其三重特征值, -3 为其单特征值, 因之经直交变换后, 它的典型式为

$$y_1^2+y_2^2+y_3^2-3y_4^2.$$

为要求出直交变换所对应的直交矩阵, 先求对应于特征值 $\lambda=1$ 的特征向量, 因而需解齐次方程组

$$\begin{cases} -\xi_1+\xi_2+\xi_3-\xi_4=0, \\ \xi_1-\xi_2-\xi_3+\xi_4=0, \\ \xi_1-\xi_2-\xi_3+\xi_4=0, \\ -\xi_1+\xi_2+\xi_3-\xi_4=0, \end{cases}$$

求得其基础解系

$$\begin{cases} Y_1=(1,\ 1,\ 0,\ 0), \\ Y_2=(1,\ 0,\ 1,\ 0), \\ Y_3=(-1,\ 0,\ 0,\ 1). \end{cases}$$

将它们先直交化, 得到

$$\begin{cases} Z_1 = Y_1 = (1,\ 1,\ 0,\ 0), \\ Z_2 = Y_2 - \dfrac{(Y_2,\ Z_1)}{(Z_1,\ Z_1)} \cdot Z_1 = \left(\dfrac{1}{2},\ -\dfrac{1}{2},\ 1,\ 0\right), \\ Z_3 = Y_3 - \dfrac{(Y_3,\ Z_1)}{(Z_1,\ Z_1)} \cdot Z_1 - \dfrac{(Y_3,\ Z_2)}{(Z_2,\ Z_2)} \cdot Z_2 \\ \quad = \left(-\dfrac{1}{3},\ \dfrac{1}{3},\ \dfrac{1}{3},\ 1\right); \end{cases}$$

再使 Z_1, Z_2, Z_3 标准化，即令 $X_i = Z_i / \|Z_i\|$ $(i = 1,\ 2,\ 3)$，就得到

$$\begin{cases} X_1 = \left(\dfrac{1}{\sqrt{2}},\ \dfrac{1}{\sqrt{2}},\ 0,\ 0\right), \\ X_2 = \left(\dfrac{1}{\sqrt{6}},\ -\dfrac{1}{\sqrt{6}},\ \dfrac{\sqrt{2}}{\sqrt{3}},\ 0\right), \\ X_3 = \left(-\dfrac{1}{2\sqrt{3}},\ \dfrac{1}{2\sqrt{3}},\ \dfrac{1}{2\sqrt{3}},\ \dfrac{\sqrt{3}}{2}\right). \end{cases}$$

其次，求对应于单特征值 $\lambda = -3$ 之特征向量，需解齐次方程组

$$\begin{cases} 3\xi_1 + \xi_2 + \xi_3 - \xi_4 = 0, \\ \xi_1 + 3\xi_2 - \xi_3 + \xi_4 = 0, \\ \xi_1 - \xi_2 + 3\xi_3 + \xi_4 = 0, \\ -\xi_1 + \xi_2 + \xi_3 + 3\xi_4 = 0, \end{cases}$$

它只有一个线性无关解向量，可选为 $Z_4 = (1,\ -1,\ -1,\ 1)$，再将它变为单位向量得

$$X_4 = \left(\frac{1}{2},\ -\frac{1}{2},\ -\frac{1}{2},\ \frac{1}{2}\right).$$

于是，以 X_1, X_2, X_3, X_4 为列向量的实直交矩阵

$$
\boldsymbol{P}=\begin{pmatrix}
\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} & -\frac{1}{2\sqrt{3}} & \frac{1}{2} \\
\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} & \frac{1}{2\sqrt{3}} & -\frac{1}{2} \\
0 & \sqrt{\frac{2}{3}} & \frac{1}{2\sqrt{3}} & -\frac{1}{2} \\
0 & 0 & \frac{\sqrt{3}}{2} & \frac{1}{2}
\end{pmatrix},
$$

即直交变换 $\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix}=\boldsymbol{P}\begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{pmatrix}$，把所给的二次齐式化为典型式

$$
y_1^2+y_2^2+y_3^2-3y_4^2.
$$

注意：这时，矩阵 $\boldsymbol{A}$ 关于多重特征值的线性无关特征向量组的选取方法不是唯一的，所以存在很多不同的直交变换，能够化 $\varphi(x_1, x_2, x_3, x_4)$ 为典型式.

为以后引用方便，特又将实对称矩阵之一些主要性质写为下面的：

定理 5 实对称矩阵的特征值都是实数. 对应于不同特征值的特征向量是互相直交的. 每个特征值的代数重数等于它的几何重数，因而秩为 r 时就恰有 0 为它的 $n-r$ 重的特征值. 又，它的最小多项式没有重根，因之其特征矩阵之初等因子都是一次式.

练习三十三

1. 试化二次齐式 $x^2+5y^2-z^2+4\sqrt{2}zx$ 到主轴上去，并求出变到主轴上的一个直交变换.

2. 设 A 是二次齐式 $\varphi(x_1, \cdots, x_n)$ 之矩阵，λ 是 A 的一个特征值。证明：一定有不完全是零的一组实数 $x_1, x_2, \cdots, x_n$ 满足方程 $\varphi(x_1, \cdots, x_n)=\lambda(x_1^2+\cdots+x_n^2)$.

3. 设有二次齐式 $\sum a_{ij}x_ix_j=X'AX$，而 λ 与 μ 分别为矩阵 A 之最大和最小的特征值。证明：对任何一组实数 $x_1, \cdots, x_n$ 常有 $\lambda X'X \geqslant X'AX \geqslant \mu X'X$.

4. 设 A 为 n 级实对称矩阵。证明：不论 $x=(x_1, x_2, \cdots, x_n)$ 与 $y=(y_1, y_2, \cdots, y_n)$ 是任何两个 n 实数组，常有 $(xA, y)=(x, yA)$.

5. 求 $A=\begin{pmatrix} 1 & 2 & 3 & \cdots & n \\ 2 & 3 & 4 & \cdots & n+1 \\ 3 & 4 & 5 & \cdots & n+2 \\ \vdots & \vdots & \vdots & & \vdots \\ n & n+1 & n+2 & \cdots & 2n-1 \end{pmatrix}$ 之最小多项式。

第5节　二次曲面(曲线)的度度分类

今利用前节主轴问题的理论，来讨论二次曲面（曲线）的度量分类。设有一个二次曲面

$$\varphi(x, y, z)\equiv ax^2+by^2+cz^2+2dxy+2exz+2fyz+2gx+2hy+2kz+l=0, \qquad (14)$$

将它写成矩阵的表示式

$$\varphi(x, y, z)\equiv(x, y, z, 1)\cdot A\cdot\begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix},$$

其中 $A=\begin{pmatrix} a & d & e & g \\ d & b & f & h \\ e & f & c & k \\ g & h & k & l \end{pmatrix}$ 是实对称矩阵，叫做 $\varphi(x, y, z)$ 的矩阵，A 之秩与行列式 $|A|$ 分别叫做 $\varphi(x, y, z)$ 之秩与行列式。首先，有下面的

引理 直角坐标轴系的变换不会使 $\varphi(x, y, z)$ 的秩与行列式发生变化.

证明 我们知道，从旧坐标系变到新坐标系（直角的）的变换公式是

$$x = l_1x' + m_1y' + n_1z' + p_1,$$
$$y = l_2x' + m_2y' + n_2z' + p_2,$$
$$z = l_3x' + m_3y' + n_3z' + p_3,$$

式中 (p_1, p_2, p_3) 是新原点对旧坐标系的坐标；而 $l_1, l_2, l_3; m_1, m_2, m_3; n_1, n_2, n_3$ 分别表示新 x' 轴，新 y' 轴，新 z' 轴对旧坐标系各轴的方向余弦. 上面的坐标变换公式可用矩阵形式来表示：

$$\begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix} = \begin{pmatrix} l_1 & m_1 & n_1 & p_1 \\ l_2 & m_2 & n_2 & p_2 \\ l_3 & m_3 & n_3 & p_3 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x' \\ y' \\ z' \\ 1 \end{pmatrix} = Q \begin{pmatrix} x' \\ y' \\ z' \\ 1 \end{pmatrix},$$

其中 $\begin{pmatrix} l_1 & m_1 & n_1 \\ l_2 & m_2 & n_2 \\ l_3 & m_3 & n_3 \end{pmatrix}$ 为直交矩阵. 这样，

$$\varphi(x, y, z) = (x, y, z, 1) \cdot A \cdot \begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix}$$

$$= (x', y', z', 1) Q' A Q \begin{pmatrix} x' \\ y' \\ z' \\ 1 \end{pmatrix},$$

说明了经坐标轴的变换后，$\varphi(x, y, z) \equiv \psi(x', y', z')$ 的矩阵为 $\boldsymbol{Q}'\boldsymbol{A}\boldsymbol{Q}$，由于 $\boldsymbol{Q}$ 之满秩性，知 $\boldsymbol{Q}'\boldsymbol{A}\boldsymbol{Q}$ 的秩等于 $\boldsymbol{A}$ 之秩，又行列式

$$|\boldsymbol{Q}'\boldsymbol{A}\boldsymbol{Q}| = |\boldsymbol{Q}'| \cdot |\boldsymbol{A}| \cdot |\boldsymbol{Q}| = |\boldsymbol{A}| \text{（因为 } |\boldsymbol{Q}| = 1\text{）}.$$

故引理成立.

再回到二次曲面(14)的研究. 由主轴问题知道，有一个直交变换(它对应着坐标原点不动的坐标轴的旋转)，使

$$ax^2 + by^2 + cz^2 + 2dxy + 2ezx + 2fyz$$

变成典型式

$$\lambda_1 x^2 + \lambda_2 y^2 + \lambda_3 z^2,$$

这时仍用 x, y, z 而不用 x', y', z' 是为了减少记号上的变更，但 $\lambda_1, \lambda_2, \lambda_3$ 是对称矩阵

$$\boldsymbol{B} = \begin{pmatrix} a & d & e \\ d & b & f \\ e & f & c \end{pmatrix}$$

的(实)特征值. 注意 $3 \leqslant r_B \leqslant r_A \leqslant 4$.

坐标轴的这样旋转变换将 $\varphi(x, y, z)$ 变成

$$\lambda_1 x^2 + \lambda_2 y^2 + \lambda_3 z^2 + 2g_1 x + 2h_1 y + 2k_1 z + l' \tag{15}$$

的形式. 这里 $l' = l$(这是因为，原点不动的旋转变换不改变常数项). 下面，分三种情况加以讨论:

① $\boldsymbol{B}$ 的秩 $r_B = 3$.

这时，三个特征值 $\lambda_1, \lambda_2, \lambda_3$ 都不为零. 所以，再平移坐标轴使新原点为 $\left(-\dfrac{g_1}{\lambda_1}, -\dfrac{h_1}{\lambda_2}, -\dfrac{k_1}{\lambda_3}\right)$，则(15)简化为

$$\lambda_1 x^2 + \lambda_2 y^2 + \lambda_3 z^2 + l''. \tag{16}$$

再细分 $r_A = 4$ 与 $r_A = 3$ 两种情况. 当 $r_A = 4$ 时，由引理得知

(16)中 $l''\neq 0$, 且 $|\boldsymbol{A}|=\lambda_1\lambda_2\lambda_3 l''$. 因而 $l''=\dfrac{|\boldsymbol{A}|}{\lambda_1\lambda_2\lambda_3}=\dfrac{|\boldsymbol{A}|}{|\boldsymbol{B}|}$, 这表示着坐标轴的适当变换可使方程(14)简化为

$$\lambda_1x^2+\lambda_2y^2+\lambda_3z^2+\frac{|\boldsymbol{A}|}{|\boldsymbol{B}|}=0,$$

此时曲面是一个有中心的曲面. 实际上, 有:

i) 若 λ_1, λ_2, λ_3 都和 $\mu=\dfrac{|\boldsymbol{A}|}{|\boldsymbol{B}|}$ 异号, 曲面(14)表椭球面;

ii) 若 λ_1, λ_2, λ_3 都和 $\mu=\dfrac{|\boldsymbol{A}|}{|\boldsymbol{B}|}$ 同号, 曲面(14)叫虚椭球面;

iii) 若 λ_1, λ_2, λ_3 中有二个与 μ 异号(一个与 μ 同号), 曲面(14)表示单叶双曲面;

iv) 若 λ_1, λ_2, λ_3 中有一个与 μ 异号(即二个与 μ 同号), 曲面(14)是双叶双曲面.

如果 $r_A=3$, 这时(16)中 $l''=0$, 方程(14)简化为

$$\lambda_1x^2+\lambda_2y^2+\lambda_3z^2=0$$

形, 这时曲面为一个锥面, 实际上是:

v) 若 λ_1, λ_2, λ_3 同号, 曲面退化, 只有一点;

vi) 若 λ_1, λ_2, λ_3 不全同号, 曲面为椭圆锥面.

② $\boldsymbol{B}$ 的秩 $r_B=2$.

这时, 三个特征值中有一为零, 另二个不为零. 不失一般性, 可设 $\lambda_3=0$, $\lambda_1\lambda_2\neq 0$. 再平移坐标轴使新原点为 $\left(-\dfrac{g_1}{\lambda_1}, -\dfrac{h_1}{\lambda_2}, 0\right)$, 则(15)简化为

$$\lambda_1x^2+\lambda_2y^2+2k_1z+l''. \tag{17}$$

它的行列式为

$$\begin{vmatrix} \lambda_1 & 0 & 0 & 0 \\ 0 & \lambda_2 & 0 & 0 \\ 0 & 0 & 0 & k_1 \\ 0 & 0 & k_1 & l'' \end{vmatrix} = -\lambda_1\lambda_2 k_1^2.$$

由引理得知 $|\boldsymbol{A}| = -\lambda_1\lambda_2 k_1^2$. 再细分如下三种情况: $r_A=4$, $r_A=3$, $r_A=2$.

在 $r_A=4$ 时, $k_1\neq 0$, 因而又平移坐标轴到新原点 $\left(0, 0, -\frac{l''}{2k_1}\right)$, 则(17)式简化为

$$\lambda_1 x^2+\lambda_2 y^2+2k_1 z,$$

故得:

vii) 当 λ_1 与 λ_2 同号时, 曲面为椭圆抛物面;

viii) 当 λ_1 与 λ_2 异号时, 曲面为双曲抛物面.

在 $r_A=3$ 时, $k_1=0$ 且必 $l''\neq 0$, 故(17)式为

$$\lambda_1 x^2+\lambda_2 y^2+l'',$$

而有:

ix) 当 λ_1 与 λ_2 异号时, 曲面为双曲柱面;

x) 当 λ_1 与 λ_2 同号时, 曲面为椭圆柱面(或实的, 在 $\lambda_1\lambda_2 l''<0$ 时; 或虚的, 在 λ_1 与 λ_2 都和 l'' 同号时).

在 $r_A=2$ 时, $k_1=l''=0$, (17)简化为

$$\lambda_1 x^2+\lambda_2 y^2,$$

于是到得:

xi) 当 $\lambda_1\lambda_2<0$ 时, 曲面为两个相交的平面;

xii) 当 $\lambda_1\lambda_2>0$ 时, 曲面退化为一条直线(新 z 轴), 或叫做相交于实直线的两个虚平面.

③ $\boldsymbol{B}$ 的秩 $r_B=1$.

这时, 有二个特征值为 0, 不失普遍性, 可令 $\lambda_1\neq 0$, $\lambda_2=$

$\lambda_3=0$. 再平移轴使新原点为$\left(-\frac{g_1}{\lambda_1},\ 0,\ 0\right)$, 得使(15)简化为

$$\lambda_1x^2+2h_1y+2k_1z+l'', \tag{18}$$

其行列式为 $$\begin{vmatrix} \lambda_1 & 0 & 0 & 0 \\ 0 & 0 & 0 & h_1 \\ 0 & 0 & 0 & k_1 \\ 0 & h_1 & k_1 & l'' \end{vmatrix}=0,$$

故由引理得知 $r_A\leqslant 3$. 因而 $r_A=3,\ 2$, 或 1.

在 $r_A=3$ 时, h_1 与 k_1 中至少有一个不为 0, 设 $h_1\neq 0$ (不失普遍性), 再平移坐标轴到新原点 $\left(0,\ -\frac{l''}{2h_1},\ 0\right)$ 可使(18)简化为

$$\lambda_1x^2+2h_1y+2k_1z.$$

再经过旋转变换

$$\begin{cases} x=x', \\ y=\dfrac{h_1}{\sqrt{h_1^2+k_1^2}}y'-\dfrac{k_1}{\sqrt{h_1^2+k_1^2}}z', \\ z=\dfrac{k_1}{\sqrt{h_1^2+k_1^2}}y'+\dfrac{h_1}{\sqrt{h_1^2+k_1^2}}z', \end{cases}$$

使上式随而方程(14)得简化为

$$\lambda_1x'^2+2\sqrt{h_1^2+k_1^2}\,y'=0,$$

由于$\sqrt{h_1^2+k_1^2}\neq 0$ 知有:

xiii) 曲面为抛物柱面.

在 $r_A=2$ 时, 必 $h_1=k_1=0$, $l''\neq 0$, 于是(18)即曲面(14)之方程简化为

$$\lambda_1x^2+l''=0.$$

故得

xiv) 曲面为二个平行的平面(实的, 在 $\lambda_1l''<0$ 时; 或虚

的，在 $\lambda_1 l''>0$ 时）。

在 $r_A=1$ 时，(18)中 h_1, k_1, l'' 必全为 0，故(14)简化为 $\lambda_1 x^2=0$，故又得

xv） 曲面是二个重合的平面。

总上所述，得到二次曲面(14)的度量分类表为

<table>
<tr><th>B 的秩</th><th>A 的秩</th><th>曲　　面</th><th colspan="2">有附带条件的细分类</th></tr>
<tr><td rowspan="4">3</td><td rowspan="4">4</td><td rowspan="4">有唯一个中心的曲面</td><td colspan="2">椭球面，当 λ_1, λ_2, λ_3 都和 $\frac{|A|}{|B|}$ 异号时</td></tr>
<tr><td colspan="2">虚椭球面，当 λ_1, λ_2, λ_3 都和 $\frac{|A|}{|B|}$ 同号时</td></tr>
<tr><td colspan="2">单叶双曲面，当 λ_1, λ_2, λ_3 中只一个与 $\frac{|A|}{|B|}$ 同号时</td></tr>
<tr><td colspan="2">双叶双曲面，当 λ_1, λ_2, λ_3 中恰有二个与 $\frac{|A|}{|B|}$ 同号时</td></tr>
<tr><td rowspan="2">3</td><td rowspan="2">3</td><td rowspan="2">锥　　面</td><td colspan="2">退化锥面(即点锥面)，λ_1, λ_2, λ_3 同号时</td></tr>
<tr><td colspan="2">非退化锥面，当 λ_1, λ_2, λ_3 异号时</td></tr>
<tr><td rowspan="2">2</td><td rowspan="2">4</td><td rowspan="2">抛物面</td><td>椭圆抛物面，当 λ_1 与 λ_2 同号时</td><td rowspan="6">$\lambda_3=0$</td></tr>
<tr><td>双曲抛物面，当 λ_1 与 λ_2 异号时</td></tr>
<tr><td rowspan="2">2</td><td rowspan="2">3</td><td rowspan="2">中心为一条直线的柱面</td><td>双曲柱面，当 λ_1 与 λ_2 异号时</td></tr>
<tr><td>椭圆柱面(实或虚)，当 λ_1 与 λ_2 同号时</td></tr>
<tr><td rowspan="2">2</td><td rowspan="2">2</td><td rowspan="2">相交的二个平面</td><td>实的，当 λ_1 与 λ_2 异号时</td></tr>
<tr><td>虚的(即退化为一直线)，当 λ_1 与 λ_2 同号时</td></tr>
<tr><td>1</td><td>4</td><td>不可能出现</td><td colspan="2"></td></tr>
<tr><td>1</td><td>3</td><td>抛物柱面</td><td colspan="2"></td></tr>
<tr><td>1</td><td>2</td><td>平行的两个平面(实或虚)</td><td colspan="2"></td></tr>
<tr><td>1</td><td>1</td><td>一个平面(即二个重合的平面)</td><td colspan="2"></td></tr>
</table>

表中 $\boldsymbol{A}=\begin{pmatrix} a & d & e & g \\ d & b & f & h \\ e & f & c & k \\ g & h & k & l \end{pmatrix}$, $\boldsymbol{B}=\begin{pmatrix} a & d & e \\ d & b & f \\ e & f & c \end{pmatrix}$, 而 $\lambda_1, \lambda_2, \lambda_3$ 为 $\boldsymbol{B}$ 之三个特征值(故为实数).

注意: 上表不是说二次曲面的度量类别只有八种. 所谓两个二次曲面属于度量的同一值类, 指的是它们可由欧氏空间之直射变换(即平移、旋转与反射的合并)互相推得. 因之, 度量类别之个数是无限的, 例如同是两个椭球面可属于不相同的度量类别, 即每类不能由秩及正负号差来决定(它们只能决定仿射与投影分类), 应由二次曲面系数间的一些关系式(叫代数不变量)来决定. 我们不谈这些东西, 因为它们属几何学范围, 象上面讲法已足说明利用代数理论解决了二次曲面的一切类型.

完全用同样方法, 可得知平面中二次曲线方程

$$ax^2+2bxy+cy^2+2dx+2ey+f=0 \tag{19}$$

的分类表为

<table>
<tr><th>$\boldsymbol{B}$ 的秩</th><th>$\boldsymbol{A}$的秩</th><th colspan="3">曲　　　　线</th></tr>
<tr><td rowspan="3">2</td><td rowspan="3">3</td><td>$|\boldsymbol{B}|<0$</td><td colspan="2">$|\boldsymbol{B}|>0$</td></tr>
<tr><td rowspan="2">双曲线</td><td>$(a+c)\cdot|\boldsymbol{A}|<0$</td><td>$(a+c)\cdot|\boldsymbol{A}|>0$</td></tr>
<tr><td>椭　圆</td><td>虚椭圆</td></tr>
<tr><td>2</td><td>2</td><td>相交直线</td><td colspan="2">一点(点椭圆)</td></tr>
<tr><td>1</td><td>3</td><td colspan="3">抛　物　线</td></tr>
<tr><td>1</td><td>2</td><td colspan="3">二平行直线(实或虚)</td></tr>
<tr><td>1</td><td>1</td><td colspan="3">二重合直线</td></tr>
</table>

表中 $\boldsymbol{A}=\begin{pmatrix} a & b & d \\ b & c & e \\ d & e & f \end{pmatrix}$, $\boldsymbol{B}=\begin{pmatrix} a & b \\ b & c \end{pmatrix}$.

至于此表的详细推导,留给读者作练习题.

本章开头曾说过本章是以主轴问题为中心问题,因此对直交矩阵应作深入的讨论,但为了方便,关于直交矩阵之性质以及它在几何上的应用等,放在下章讲酉矩阵时再讲,因此本节到此结束.

练 习 三 十 四

1. 详细推导上面的二次曲线的分类表中各种情况.
2. 试用本节的方法,判定下列曲面的类型:

i) $x^2+3y^2+5z^2+4xy+6xz+8yz+8x+10y+12z+7=0$,

ii) $x^2+y^2-3z^2+4xy-4xz+4yz-2x-2y-2z+1=0$.

第 6 节 欧氏空间中的几个零星问题

在平面 $\Re_2$ 中(图 7–6),已知过原点 O 之任一直线 l 为一维子空间(向量空间,向量的起点放在原点). 又知有唯一个过原点之直线 $l'\perp l$, 即有唯一个一维向量子空间 l' 与 l 互相垂直. 很显然, l 与 l' 的和 $l\cup l'=\Re_2$ (即 $\Re_2$ 中任一向量可表成 l 中一向量与 l' 中一向量之和).

又如图 7–7, 在三维空间 $\Re_3$ 中过原点 O 的一维向量子空间 l 也有唯一个平面 π(即二维向量子空间 π)$\perp l$. 在图 7–8 中, $\Re_3$ 之二维向量子空间 π(即平面 π)在 $\Re_3$ 内也有唯一个一维子空间 l (即过原点 O 之一直线 l)$\perp\pi$. 且不论图 7-7 或图 7–8, 恒有 $\Re_3=\pi\cup l$.

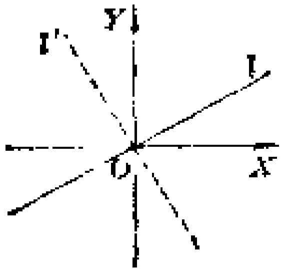


图 7-6

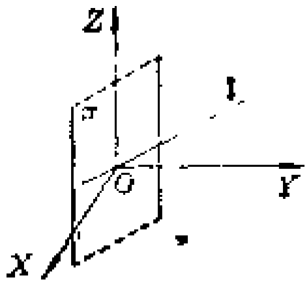


图 7-7

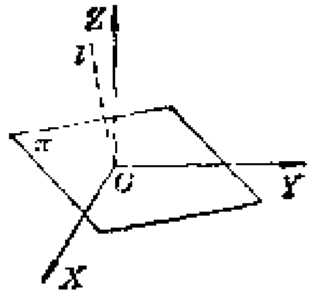


图 7-8

由上述关于 $\mathfrak{R}_2$ 与 $\mathfrak{R}_3$ 之启示，今问 n 维欧氏空间 $\mathfrak{R}_n$ 中是否也有类似的结果呢？首先要明了欧氏空间 $\mathfrak{R}_n$ 中两个子空间直交(即垂直)之意义.

定义 1 欧氏空间 $\mathfrak{R}_n$ 中两个子空间 $\mathfrak{V}_1$ 与 $\mathfrak{V}_2$ 若有这样的性质，即对任意的 $\boldsymbol{\alpha}_1\in\mathfrak{V}_1$ 与任意的 $\boldsymbol{\alpha}_2\in\mathfrak{V}_2$ 恒有 $(\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2)=0$ 时，就叫 $\mathfrak{V}_1$ 与 $\mathfrak{V}_2$ 是直交的，用符号 $\mathfrak{V}_1\perp\mathfrak{V}_2$ 来表示.

当某向量 $\boldsymbol{\alpha}$ 与子空间 $\mathfrak{V}$ 之任何向量 $\boldsymbol{x}$(任 $\boldsymbol{x}\in\mathfrak{V}$) 都直交，即 $(\boldsymbol{\alpha}, \boldsymbol{x})=0$ 时，就叫向量 $\boldsymbol{\alpha}$ 与子空间 $\mathfrak{V}$ 直交，记为 $\boldsymbol{\alpha}\perp\mathfrak{V}$.

显然，由于 $(\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2)=(\boldsymbol{\alpha}_2, \boldsymbol{\alpha}_1)$，故当 $\mathfrak{V}_1\perp\mathfrak{V}_2$ 时也有 $\mathfrak{V}_2\perp\mathfrak{V}_1$.

又因为只有零向量 $\mathbf{0}$ 才与它自身直交，故从 $\mathfrak{V}_1\perp\mathfrak{V}_2$ 马上知道 $\mathfrak{V}_1\cap\mathfrak{V}_2=\mathfrak{O}$.

关于直交的子空间，有下面的

定理 1 *如果子空间 $\mathfrak{V}_1, \mathfrak{V}_2, \cdots, \mathfrak{V}_s$ 两两直交，即 $\mathfrak{V}_i\perp\mathfrak{V}_j$ $(i\neq j)$，那末和 $\mathfrak{V}_1\cup\mathfrak{V}_2\cup\cdots\cup\mathfrak{V}_s$ 为直和，记为 $\mathfrak{V}_1\oplus\mathfrak{V}_2\oplus\cdots\oplus\mathfrak{V}_s$，即 $\mathfrak{V}_1\cup\mathfrak{V}_2\cup\cdots\cup\mathfrak{V}_s$ 之每向量 $\boldsymbol{x}=\boldsymbol{x}_1+\boldsymbol{x}_2+\cdots+\boldsymbol{x}_s$ $(\boldsymbol{x}_i\in\mathfrak{V}_i)$ 的表示法是唯一的.*

证明 若 $\boldsymbol{x}_i\in\mathfrak{V}_i$ $(i=1, 2, \cdots, s)$ 且 $\boldsymbol{x}_1+\cdots+\boldsymbol{x}_s=\mathbf{0}$，则

$$0=(\mathbf{0},\ \boldsymbol{x}_j)=\left(\sum_{i=1}^{s}\boldsymbol{x}_i,\ \boldsymbol{x}_j\right)=\sum_{i=1}^{s}(\boldsymbol{x}_i,\ \boldsymbol{x}_j)=(\boldsymbol{x}_j,\ \boldsymbol{x}_j),$$

所以 $\boldsymbol{x}_j=\mathbf{0}(j=1,\ 2,\ \cdots,\ s)$. 这说明了零向量 $\mathbf{0}$ 写为 $\boldsymbol{x}_1+\cdots+\boldsymbol{x}_s$ 之表示法($\boldsymbol{x}_i\in\mathfrak{V}_i$)是唯一的，即只能是每 $\boldsymbol{x}_i=\mathbf{0}$，因而 $\mathfrak{V}_1\cup\cdots\cup\mathfrak{V}_s$ 之每向量 $\boldsymbol{x}=\sum_{i=1}^{s}\boldsymbol{x}_i$ 之表示法亦唯一，也就是有直和 $\mathfrak{V}_1\oplus\cdots\oplus\mathfrak{V}_s$. 证完.

定理 2 欧氏空间 $\mathfrak{R}_n$ 中对于每一个子空间 $\mathfrak{V}$ 都有唯一一个子空间 $\mathfrak{U}$ 使 $\mathfrak{V}\perp\mathfrak{U}$，且 $\mathfrak{R}_n=\mathfrak{V}\oplus\mathfrak{U}$. ——称 $\mathfrak{U}$ 为 $\mathfrak{V}$ 之直交补，记为 $\mathfrak{U}=\mathfrak{V}^{\perp}$，因之 $\mathfrak{V}=\mathfrak{U}^{\perp}$.

证明 若 $\mathfrak{V}=\mathbf{0}$，则显有 $\mathfrak{R}_n=\mathfrak{V}^{\perp}$，唯一性显然. 设 $\mathfrak{V}$ 为非零子空间，由于在 $\mathfrak{R}_n$ 中所定义的内积之意义也得使 $\mathfrak{V}$ 成为欧氏空间，故若 $\mathfrak{V}$ 之维数等于 m，则据第 2 节定理 2 可知 $\mathfrak{V}$ 有标准直交基底，今令 $\boldsymbol{e}_1,\ \boldsymbol{e}_2,\ \cdots,\ \boldsymbol{e}_m$ 为 $\mathfrak{V}$ 的一组标准直交基底. 如果 $m=n$，则 $\mathfrak{V}=\mathfrak{R}_n$，因之 $\mathfrak{V}^{\perp}=\mathbf{0}$ 且 $\mathfrak{V}^{\perp}$ 之唯一性也显然. 故需讨论的是 $0<m<n$. 这时，必有 $\boldsymbol{\alpha}\in\mathfrak{R}_n$ 使 $\boldsymbol{\alpha}\notin\mathfrak{V}$，因而 $\boldsymbol{\alpha}$ 不能为 $\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_m$ 的线性组合，再作向量 $\boldsymbol{e}_{m+1}=\boldsymbol{\alpha}+\lambda_1\boldsymbol{e}_1+\cdots+\lambda_m\boldsymbol{e}_m$(其中 $\lambda_1,\lambda_2,\cdots,\lambda_m$ 是待定的常数)，于是利用 $\boldsymbol{e}_1,\ \cdots,\ \boldsymbol{e}_m$ 间的标准直交性得知

$$\begin{aligned}(\boldsymbol{e}_i,\ \boldsymbol{e}_{m+1})&=(\boldsymbol{e}_i,\ \boldsymbol{\alpha})+\sum_{j=1}^{m}\lambda_j\cdot(\boldsymbol{e}_i,\boldsymbol{e}_j)\\&=(\boldsymbol{e}_i,\boldsymbol{\alpha})+\lambda_i\qquad(i=1,\ 2,\ \cdots,\ m).\end{aligned}$$

故取 $\lambda_i=-(\boldsymbol{e}_i,\ \boldsymbol{\alpha})$时就有$(\boldsymbol{e}_i,\ \boldsymbol{e}_{m+1})=0$，因而由 $\boldsymbol{e}_{m+1}$ 之形成方法又知 $\boldsymbol{e}_{m+1}\neq\mathbf{0}$，于是又得令 $\boldsymbol{e}_{m+1}$ 标准化，即可令 $(\boldsymbol{e}_{m+1},\ \boldsymbol{e}_{m+1})=1$. 这说明了 $m<n$ 时，由 $\mathfrak{V}$ 之一组标准直交基底 $\boldsymbol{e}_1,\ \cdots,\ \boldsymbol{e}_m$ 可扩充为 $m+1$ 个两两直交且长为 1 的一组向量. 若 $m+1<n$，又可继续这样的方法往下做，一直到 $m+k=n$ 为止. 这无异乎是说由欧氏空间 $\mathfrak{R}_n$ 之子空间 $\mathfrak{V}$ 的一组标准直

交基底可扩充成为 $\mathfrak{R}_n$ 之一组标准直交基底†. 于是, 在 $\mathfrak{R}_n$ 中必可找得 $n-m$ 个向量 $\boldsymbol{e}_{m+1}, \cdots, \boldsymbol{e}_n$ 使

$$\boldsymbol{e}_1, \boldsymbol{e}_2, \cdots, \boldsymbol{e}_m, \boldsymbol{e}_{m+1}, \cdots, \boldsymbol{e}_n$$

为 $\mathfrak{R}_n$ 之一组标准直交基底. 再作 $\mathfrak{R}_n$ 之子空间 $\mathfrak{U}=\mathfrak{S}[\boldsymbol{e}_{m+1}, \cdots, \boldsymbol{e}_n]$, 即 $\mathfrak{U}$ 由 $\boldsymbol{e}_{m+1}, \cdots, \boldsymbol{e}_n$ 所生成, 显然有 $\mathfrak{R}_n=\mathfrak{V}\oplus\mathfrak{U}$ 且 $\mathfrak{U}\perp\mathfrak{V}$. 这说明了 $\mathfrak{U}$ 为 $\mathfrak{V}$ 之一个直交补.

再证唯一性. 若尚有 $\mathfrak{R}_n=\mathfrak{V}\oplus\mathfrak{W}$ 且 $\mathfrak{W}\perp\mathfrak{V}$, 则当 $\boldsymbol{a}\in\mathfrak{W}$ 时, 从关系式 $\boldsymbol{a}\in\mathfrak{R}_n=\mathfrak{V}\oplus\mathfrak{U}$ 可知 $\boldsymbol{a}=\boldsymbol{v}+\boldsymbol{u}$ (这里 $\boldsymbol{v}\in\mathfrak{V}$, $\boldsymbol{u}\in\mathfrak{U}$); 因之利用 $\mathfrak{W}\perp\mathfrak{V}$ 可知 $0=(\boldsymbol{a}, \boldsymbol{v})=(\boldsymbol{v}+\boldsymbol{u}, \boldsymbol{v})=(\boldsymbol{v}, \boldsymbol{v})+(\boldsymbol{u}, \boldsymbol{v})=(\boldsymbol{v}, \boldsymbol{v})$ (其中, 利用了 $(\boldsymbol{u}, \boldsymbol{v})=0$), 所以 $\boldsymbol{v}=\boldsymbol{0}$; 故 $\boldsymbol{a}=\boldsymbol{u}\in\mathfrak{U}$, 即证明了 $\mathfrak{W}\subseteq\mathfrak{U}$. 同样的道理, 可知 $\mathfrak{U}\subseteq\mathfrak{W}$. 所以 $\mathfrak{W}=\mathfrak{U}$, 唯一性获证.

定理 2 证完.

由定理 2 之证明方法又得到下面的

推论 欧氏空间 $\mathfrak{R}_n$ 之子空间 $\mathfrak{V}$ 的直交补 $\mathfrak{V}^{\perp}$ 恰由一切与 $\mathfrak{V}$ 成直交的向量组成, 且 $\mathfrak{V}$ 之维数 $+\mathfrak{V}^{\perp}$ 之维数 $=n$.

下面再来讨论欧氏空间 $\mathfrak{R}_n$ 中线性变换对于子空间 $\mathfrak{V}$ 及其直交补 $\mathfrak{V}^{\perp}$ 的影响怎样?

为研究的方便, 取 $\boldsymbol{e}_1, \boldsymbol{e}_2, \cdots, \boldsymbol{e}_n$ 为 $\mathfrak{R}_n$ 的一组标准直交基底; 如上所述, 可设 $\boldsymbol{e}_1, \cdots, \boldsymbol{e}_m$ 为 $\mathfrak{V}$ 之基底, 而 $\boldsymbol{e}_{m+1}, \cdots, \boldsymbol{e}_n$ 为 $\mathfrak{V}^{\perp}$ 之基底.

取 $\mathfrak{R}_n$ 之任一个线性变换 $\mathscr{T}$, 并令 $\boldsymbol{e}_i\mathscr{T}=\sum\limits_{s=1}^{n} a_{is}\boldsymbol{e}_s$ $(i=1,$

---

† 这一结果很重要, 特作为定理于下:

定理 n 维欧氏空间 $\mathfrak{R}_n$ 中任一个两两标准直交的向量组可扩充成 $\mathfrak{R}_n$ 的一组标准基底. 同样, $\mathfrak{R}_n$ 中任一个两两直交的向量组可扩充成 $\mathfrak{R}_n$ 之一组直交基底.

2, …, n), 即 $\mathscr{T}$ 关于基底 $\boldsymbol{e}_1, \cdots, \boldsymbol{e}_n$ 所对应之矩阵为 $\boldsymbol{A}=(a_{ij})$（今后所谓某线性变换对应的矩阵，都是指对这组基底 $\boldsymbol{e}_1, \cdots, \boldsymbol{e}_n$ 而言的）. 今作一线性变换 $\mathscr{T}^*$, 令它对应之矩阵为 $\boldsymbol{A}=(a_{ij})$ 之转置矩阵 $\boldsymbol{A}'=(a_{ji})$, 即

$$\boldsymbol{e}_i\mathscr{T}^*=\sum_{s=1}^{n} a_{is}\boldsymbol{e}_s \quad (i=1, 2, \cdots, n),$$

而叫 $\mathscr{T}^*$ 为 $\mathscr{T}$ 之共轭变换.

取 $\Re_n$ 之任二个向量 $\boldsymbol{x}=\sum_{i=1}^{n}x_i\boldsymbol{e}_i$, $\boldsymbol{y}=\sum_{j=1}^{n}y_j\boldsymbol{e}_j$, 于是, 利用 $\boldsymbol{e}_i$ 间之标准直交性, 就有

$$\begin{aligned}(\boldsymbol{x}, \boldsymbol{y}\mathscr{T})&=\left(\sum_{i=1}^{n}x_i\boldsymbol{e}_i, \sum_{j=1}^{n}y_j(\boldsymbol{e}_j\mathscr{T})\right)\\&=\sum_{i,j=1}^{n}x_iy_j\cdot(\boldsymbol{e}_i, \boldsymbol{e}_j\mathscr{T})\\&=\sum_{i,j=1}^{n}x_iy_j\cdot\sum_{s=1}^{n}a_{js}\cdot(\boldsymbol{e}_i, \boldsymbol{e}_s)=\sum_{i,j=1}^{n}x_iy_ja_{ji},\end{aligned}$$

以及

$$\begin{aligned}(\boldsymbol{x}\mathscr{T}^*, \boldsymbol{y})&=\left(\sum_{i=1}^{n}x_i(\boldsymbol{e}_i\mathscr{T}^*), \sum_{j=1}^{n}y_j\boldsymbol{e}_j\right)\\&=\sum_{i,j=1}^{n}x_iy_j\sum_{s=1}^{n}a_{is}\cdot(\boldsymbol{e}_s, \boldsymbol{e}_j)=\sum_{i,j=1}^{n}x_iy_ja_{ji},\end{aligned}$$

说明了 $(\boldsymbol{x}, \boldsymbol{y}\mathscr{T})=(\boldsymbol{x}\mathscr{T}^*, \boldsymbol{y})$. 这是说欧氏空间 $\Re_n$ 中线性变换 $\mathscr{T}$ 与其共轭变换 $\mathscr{T}^*$ 恒满足下面的关系式

$$(\boldsymbol{x}, \boldsymbol{y}\mathscr{T})=(\boldsymbol{x}\mathscr{T}^*, \boldsymbol{y}), \tag{20}$$

其中 $\boldsymbol{x}, \boldsymbol{y}$ 为 $\Re_n$ 的任二个向量. 反之, 适合(20)之 $\mathscr{T}$ 与 $\mathscr{T}^*$ 也互为共轭的.

由共轭变换之意义, 容易证明下列诸性质 i～iv:

i. $(\mathscr{T}^*)^*=\mathscr{T}$;

ii. $(\mathscr{T}+\mathscr{S})^*=\mathscr{T}^*+\mathscr{S}^*$;

iii. $(k\mathscr{T})^*=k\mathscr{T}^*$;

iv. $(\mathscr{T}\mathscr{S})^*=\mathscr{S}^*\mathscr{T}^*$.

事实上，有下面的

引理 若 $\mathscr{T}$ 为欧氏空间 $\mathfrak{R}_n$ 的线性变换，且对任 $\boldsymbol{x}, \boldsymbol{y}\in \mathfrak{R}_n$ 恒有 $(\boldsymbol{x}\mathscr{T}, \boldsymbol{y})=0$，则必 $\mathscr{T}=\mathscr{O}$ (零变换)。

这是因为有 $(\boldsymbol{e}_i\mathscr{T}, \boldsymbol{e}_j)=0(i, j=1, \cdots, n)$，但 $\boldsymbol{e}_1, \cdots, \boldsymbol{e}_n$ 取为 $\mathfrak{R}_n$ 之一组标准直交基底，故令 $\boldsymbol{e}_i\mathscr{T}=\sum\limits_{s=1}^{n} a_{is}\boldsymbol{e}_s$ 时则有 $0=(\boldsymbol{e}_i\mathscr{T}, \boldsymbol{e}_j)=\sum\limits_{s=1}^{n} a_{is}\cdot(\boldsymbol{e}_s, \boldsymbol{e}_j)=a_{ij}$，即 $\boldsymbol{e}_i\mathscr{T}=\boldsymbol{0}(i=1, \cdots, n)$，所以 $\mathscr{T}=\mathscr{O}$，证完。

有了这引理，可证上述性质 i～iv.

因由共轭变换之意义，可知 $(\boldsymbol{x}(\mathscr{T}^*)^*, \boldsymbol{y})=(\boldsymbol{x}, \boldsymbol{y}\mathscr{T}^*)=(\boldsymbol{y}\mathscr{T}^*, \boldsymbol{x})=(\boldsymbol{y}, \boldsymbol{x}\mathscr{T})=(\boldsymbol{x}\mathscr{T}, \boldsymbol{y})$，因而 $(\boldsymbol{x}[(\mathscr{T}^*)^*-\mathscr{T}], \boldsymbol{y})=0$，故据引理得 $(\mathscr{T}^*)^*=\mathscr{T}$，即 i 成立。

又，$(\boldsymbol{x}(\mathscr{T}+\mathscr{S})^*, \boldsymbol{y})=(\boldsymbol{x}, \boldsymbol{y}(\mathscr{T}+\mathscr{S}))=(\boldsymbol{x}, \boldsymbol{y}\mathscr{T})+(\boldsymbol{x}, \boldsymbol{y}\mathscr{S})=(\boldsymbol{x}\mathscr{T}^*, \boldsymbol{y})+(\boldsymbol{x}\mathscr{S}^*, \boldsymbol{y})=(\boldsymbol{x}(\mathscr{T}^*+\mathscr{S}^*), \boldsymbol{y})$，而据引理即得 $(\mathscr{T}+\mathscr{S})^*=\mathscr{T}^*+\mathscr{S}^*$，故 ii 成立。

再者，$(\boldsymbol{x}(k\mathscr{T})^*, \boldsymbol{y})=(\boldsymbol{x}, \boldsymbol{y}(k\mathscr{T}))=(\boldsymbol{x}, k(\boldsymbol{y}\mathscr{T}))=k\cdot(\boldsymbol{x}, \boldsymbol{y}\mathscr{T})=k\cdot(\boldsymbol{x}\mathscr{T}^*, \boldsymbol{y})=(k(\boldsymbol{x}\mathscr{T}^*), \boldsymbol{y})=(\boldsymbol{x}(k\mathscr{T}^*), \boldsymbol{y})$，故 $(k\mathscr{T})^*=k\mathscr{T}^*$，即 iii 成立。

最后，$(\boldsymbol{x}(\mathscr{T}\mathscr{S})^*, \boldsymbol{y})=(\boldsymbol{x}, \boldsymbol{y}(\mathscr{T}\mathscr{S}))=(\boldsymbol{x}, (\boldsymbol{y}\mathscr{T})\mathscr{S})=(\boldsymbol{x}\mathscr{S}^*, \boldsymbol{y}\mathscr{T})=((\boldsymbol{x}\mathscr{S}^*)\mathscr{T}^*, \boldsymbol{y})=(\boldsymbol{x}\mathscr{S}^*\mathscr{T}^*, \boldsymbol{y})$ 可导致 $(\mathscr{T}\mathscr{S})^*=\mathscr{S}^*\mathscr{T}^*$，即 iv 成立。

现在可讨论线性变换及其共轭变换在欧氏空间内的影响。首先，有：

定理 3 设 $\mathfrak{R}_n$ 为欧氏空间。如果 $\mathfrak{V}$ 是线性变换 $\mathscr{T}$ 的不变子空间，那末 $\mathfrak{V}$ 的直交补 $\mathfrak{V}^{\perp}$ 是共轭变换 $\mathscr{T}^*$ 的不变子空

间.

证明 设 $\alpha\in\mathfrak{V}^{\perp}$. 于是对任 $\beta\in\mathfrak{V}$, 因有 $\beta\mathscr{T}\in\mathfrak{V}$, 故 $(\alpha, \beta\mathscr{T})=0$, 因而 $0=(\alpha, \beta\mathscr{T})=(\alpha\mathscr{T}^*, \beta)$, 所以由 β 在 $\mathfrak{V}$ 内的任意性, 故有 $\alpha\mathscr{T}^*\in\mathfrak{V}^{\perp}$. 证完.

又, 共轭变换还可以刻划直交变换, 具体说来, 有:

定理 4 欧氏空间 $\mathfrak{R}_n$ 内的线性变换 $\mathscr{T}$ 为直交的充要条件是 $\mathscr{T}\mathscr{T}^*$ 为恒等变换(即 $\mathscr{T}$ 与共轭 $\mathscr{T}^*$ 互逆).

证明 若 $\mathscr{T}$ 是直交变换, 则对任 $x, y\in\mathfrak{R}_n$ 恒有 $(x, y)=(x\mathscr{T}, y\mathscr{T})$, 故有 $(x\mathscr{E}, y)=(x, y)=(x\mathscr{T}, y\mathscr{T})=((x\mathscr{T})\mathscr{T}^*, y)=(x\mathscr{T}\mathscr{T}^*, y)$, 不得不有 $\mathscr{T}\mathscr{T}^*=\mathscr{E}$. 反之, $\mathscr{T}\mathscr{T}^*=\mathscr{E}$ 可诱致 $(x, y)=(x\mathscr{E}, y)=(x\mathscr{T}\mathscr{T}^*, y)=((x\mathscr{T})\mathscr{T}^*, y)=(x\mathscr{T}, y\mathscr{T})$, 即 $\mathscr{T}$ 是直交变换. 证完.

最后, 在欧氏空间 $\mathfrak{R}_n$ 中若线性变换 $\mathscr{T}$ 与其共轭变换 $\mathscr{T}^*$ 相等, 即 $\mathscr{T}=\mathscr{T}^*$, 就叫 $\mathscr{T}$ 为自共轭变换或对称变换. 由于对标准直交基底说来, $\mathscr{T}$ 与 $\mathscr{T}^*$ 之矩阵互为转置的, 故 $\mathscr{T}$ 为对称变换的充要条件是它的矩阵为对称矩阵.

第八章　酉　空　间

第七章里，在实线性空间中引进了"数积"概念，而使这空间成为欧氏空间；并以主轴问题为中心讨论了欧氏空间的基本理论．本章就在复线性空间里引进数积这一概念，而使之成为酉空间(也叫复欧氏空间)，并讨论酉空间的基本理论．也可以说，本章是第七章的推广，其内容除主轴问题外，还有二次齐式耦以及实直交矩阵的标准形及其几何意义．本章中有很多的概念、结论和方法都与第七章相类似，由于在第七章里已有详尽的叙述，故在这里只简单地提出，不作更多的重复．

第1节　酉空间、酉基底

在 n 维复空间 $\mathfrak{R}$ 中，若有一方法使任两个向量 $\boldsymbol{x}$, $\boldsymbol{y}$ 对应于一个确定的复数，记为 $(\boldsymbol{x}, \boldsymbol{y})$，并且适合下列的四个条件：

i. $(\boldsymbol{x}, \boldsymbol{y})=\overline{(\boldsymbol{y}, \boldsymbol{x})}$,

ii. $(\lambda\boldsymbol{x}, \boldsymbol{y})=\lambda\cdot(\boldsymbol{x}, \boldsymbol{y})$, λ 为复数,

iii. $(\boldsymbol{x}+\boldsymbol{y}, \boldsymbol{z})=(\boldsymbol{x}, \boldsymbol{z})+(\boldsymbol{y}, \boldsymbol{z})$,

iv. $(\boldsymbol{x}, \boldsymbol{x})\geqslant 0$, $(\boldsymbol{x}, \boldsymbol{x})=0$ 当且仅当 $\boldsymbol{x}=\boldsymbol{0}$ 时,

那末，叫 $(\boldsymbol{x}, \boldsymbol{y})$ 为向量 $\boldsymbol{x}$, $\boldsymbol{y}$ 的数积，而叫复空间 $\mathfrak{R}$ 为酉空间．

注意：这定义与欧氏空间中数积定义不同之处只在公理 i；因为有了 $(\boldsymbol{x}, \boldsymbol{y})=\overline{(\boldsymbol{y}, \boldsymbol{x})}$，即知 $(\boldsymbol{x}, \boldsymbol{x})$ 是实数；假如没有这个限制，$(\boldsymbol{x}, \boldsymbol{x})$ 为一复数，则公理 iv 就没有意义．同上章

一样，我们叫 $\|\boldsymbol{x}\|=\sqrt{(\boldsymbol{x}, \boldsymbol{x})}$ 为向量 $\boldsymbol{x}$ 的长．但是 $\cos(\widehat{\boldsymbol{x}, \boldsymbol{y}})$ 就不能象上章那样去定义了，只是当 $(\boldsymbol{x}, \boldsymbol{y})=0$ 时我们仍叫向量 $\boldsymbol{x}$ 与 $\boldsymbol{y}$ 直交，用记号 $\boldsymbol{x}\perp\boldsymbol{y}$ 表示之．

由这个定义，可知非零向量 $\boldsymbol{x}$ 的长 $\|\boldsymbol{x}\|$ 总是正数，只有零向量 $\mathbf{0}$ 的长 $\|\mathbf{0}\|$ 才等于零． 又由定义容易得证 $(\boldsymbol{x}, \lambda\boldsymbol{y})=\bar{\lambda}\cdot(\boldsymbol{x}, \boldsymbol{y})$ 与 $(\boldsymbol{x}, \boldsymbol{y}+\boldsymbol{z})=(\boldsymbol{x}, \boldsymbol{y})+(\boldsymbol{x}, \boldsymbol{z})$． 于是易证 $(\boldsymbol{x}, \mathbf{0})=0=(\mathbf{0}, \boldsymbol{x})$，这说明零向量与任何向量直交． 在酉空间中，两向量 $\boldsymbol{x}, \boldsymbol{y}$ 的夹角虽不能定义，但这时仍然有柯西不等式：

$$(\boldsymbol{x}, \boldsymbol{y})\cdot\overline{(\boldsymbol{x}, \boldsymbol{y})}=(\boldsymbol{x}, \boldsymbol{y})\cdot(\boldsymbol{y}, \boldsymbol{x})$$
$$\leqq\|\boldsymbol{x}\|^2\cdot\|\boldsymbol{y}\|^2=(\boldsymbol{x}, \boldsymbol{x})\cdot(\boldsymbol{y}, \boldsymbol{y}), \qquad (1)$$

证明的方法与欧氏空间的场合一样： 由于

$$(\xi\boldsymbol{x}+\eta\boldsymbol{y}, \xi\boldsymbol{x}+\eta\boldsymbol{y})=\xi\bar{\xi}\cdot(\boldsymbol{x}, \boldsymbol{x})+\xi\bar{\eta}\cdot(\boldsymbol{x}, \boldsymbol{y})$$
$$+\bar{\xi}\eta\cdot\overline{(\boldsymbol{x}, \boldsymbol{y})}+\eta\bar{\eta}\cdot(\boldsymbol{y}, \boldsymbol{y})\geqslant 0$$

是变量 ξ, η 的一个正定或半正定厄米特齐式，故

$$\begin{vmatrix}(\boldsymbol{x}, \boldsymbol{x}) & (\boldsymbol{x}, \boldsymbol{y})\\ \overline{(\boldsymbol{x}, \boldsymbol{y})} & (\boldsymbol{y}, \boldsymbol{y})\end{vmatrix}\geqslant 0,$$

即
$$(\boldsymbol{x}, \boldsymbol{y})\cdot\overline{(\boldsymbol{x}, \boldsymbol{y})}\leqslant(\boldsymbol{x}, \boldsymbol{x})\cdot(\boldsymbol{y}, \boldsymbol{y}).$$

［例］ 在由一切 n 个复数组 $\boldsymbol{x}=(x_1, x_2, \cdots, x_n)$ 所成的复空间 $\mathfrak{R}_n$ 内，如果定义两向量 $\boldsymbol{x}=(x_1, x_2, \cdots, x_n)$ 与 $\boldsymbol{y}=(y_1, y_2, \cdots, y_n)$ 的数积为 $(\boldsymbol{x}, \boldsymbol{y})=\sum_{i=1}^{n}x_i\bar{y}_i$，容易验证这时公理 i～iv 是适合的，于是 $\mathfrak{R}_n$ 变成酉空间，而柯西不等式(1)为

$$\sum_{i=1}^{n}x_i\bar{y}_i\cdot\sum_{i=1}^{n}\bar{x}_iy_i\leqslant\sum_{i=1}^{n}x_i\bar{x}_i\cdot\sum_{i=1}^{n}y_i\bar{y}_i.$$

与由实空间变成欧氏空间的情况一样，我们可以证明下列的

定理1 复空间 $\mathfrak{R}$ 成为酉空间的充要条件是两向量的数

积是由一个正定厄米特齐式来定义的.

这就是说：取定了 n 维空间 $\mathfrak{R}$ 的一组基底 $\boldsymbol{e}_1, \boldsymbol{e}_2, \cdots, \boldsymbol{e}_n$ 后，如果 $\mathfrak{R}$ 是酉空间，那末由数积 $(\boldsymbol{e}_i, \boldsymbol{e}_j)$ 所决定的 n 级矩阵

$$\begin{pmatrix} (\boldsymbol{e}_1, \boldsymbol{e}_1) & (\boldsymbol{e}_1, \boldsymbol{e}_2) & \cdots & (\boldsymbol{e}_1, \boldsymbol{e}_n) \\ (\boldsymbol{e}_2, \boldsymbol{e}_1) & (\boldsymbol{e}_2, \boldsymbol{e}_2) & \cdots & (\boldsymbol{e}_2, \boldsymbol{e}_n) \\ \cdots & \cdots & \cdots & \cdots \\ (\boldsymbol{e}_n, \boldsymbol{e}_1) & (\boldsymbol{e}_n, \boldsymbol{e}_2) & \cdots & (\boldsymbol{e}_n, \boldsymbol{e}_n) \end{pmatrix}$$

是正定厄米特矩阵；反过来，若事先给定了一个正定厄米特矩阵 $\boldsymbol{A}=(a_{ij})$，用 $(\boldsymbol{e}_i, \boldsymbol{e}_j)=a_{ij}$ 定义 $\boldsymbol{e}_i$、$\boldsymbol{e}_j$ 的数积可以使 $\mathfrak{R}$ 成为酉空间.

由是，当 $\boldsymbol{e}_1, \boldsymbol{e}_2, \cdots, \boldsymbol{e}_n$ 与 $\boldsymbol{f}_1, \boldsymbol{f}_2, \cdots, \boldsymbol{f}_n$ 是酉空间 $\mathfrak{R}$ 的两组基底时，它们所决定的正定厄米特矩阵各为

$$\boldsymbol{S}_e=\begin{pmatrix} (\boldsymbol{e}_1, \boldsymbol{e}_1) & \cdots & (\boldsymbol{e}_1, \boldsymbol{e}_n) \\ \cdots & \cdots & \cdots \\ (\boldsymbol{e}_n, \boldsymbol{e}_1) & \cdots & (\boldsymbol{e}_n, \boldsymbol{e}_n) \end{pmatrix}$$

与
$$\boldsymbol{S}_f=\begin{pmatrix} (\boldsymbol{f}_1, \boldsymbol{f}_1) & \cdots & (\boldsymbol{f}_1, \boldsymbol{f}_n) \\ \cdots & \cdots & \cdots \\ (\boldsymbol{f}_n, \boldsymbol{f}_1) & \cdots & (\boldsymbol{f}_n, \boldsymbol{f}_n) \end{pmatrix},$$

则与欧氏空间的情形一样，能证明

定理 2 $\boldsymbol{S}_f=\boldsymbol{Q}\boldsymbol{S}_e\overline{\boldsymbol{Q}}'$，这里 $\begin{pmatrix} \boldsymbol{f}_1 \\ \vdots \\ \boldsymbol{f}_n \end{pmatrix}=\boldsymbol{Q}\begin{pmatrix} \boldsymbol{e}_1 \\ \vdots \\ \boldsymbol{e}_n \end{pmatrix}$.

定理 2 表示着酉空间 $\mathfrak{R}$ 中向量数积的表达式随基底变换后的公式.

n 维酉空间 $\mathfrak{R}$ 之一组基底 $\boldsymbol{e}_1, \boldsymbol{e}_2, \cdots, \boldsymbol{e}_n$ 中诸向量若两两互为直交，就叫 $\boldsymbol{e}_1, \boldsymbol{e}_2, \cdots, \boldsymbol{e}_n$ 是 $\mathfrak{R}$ 的一组直交基底；如果每个 $\boldsymbol{e}_i$ 的长又都等于 1，即

$$(e_i, e_j)=\begin{cases}1, & 当 i=j 时,\\ 0, & 当 i\neq j 时,\end{cases}$$

就叫 $e_1, e_2, \cdots, e_n$ 是 $\mathfrak{R}$ 的一组酉基底.

完全同第七章第 2 节中所讲的一样，能证明一组两两直交的向量是线性无关的，因此在 n 维酉空间中 n 个非零的两两直交的向量就是它的直交基底.

也和欧氏空间的情形完全一样，我们可以从酉空间 $\mathfrak{R}$ 的一组基底使之直交化，就得到 $\mathfrak{R}$ 的一组直交基底；再标准化，就得到它的酉基底．于是，又有：

定理 3 n 维酉空间必有酉基底.

关于数积在酉基底内的表达式，则有：

定理 4 假设 $e_1, e_2, \cdots, e_n$ 是酉空间 $\mathfrak{R}$ 的酉基底，则任两向量 $x=\sum_1^n \xi_i e_i$, $y=\sum_1^n \eta_i e_i$ 之数积为：

$$(x, y)=\sum_1^n \xi_i\bar{\eta}_i.$$

也就是说，这时数积是由单位矩阵这一特殊的正定厄米特矩阵来定义的.

凡由一切 n 个复数组 $(x_1, x_2, \cdots, x_n)$ 构成的酉空间 $\mathfrak{R}_n$，两向量之数积的意义总是象上面那样地去定义，以求简便，这时 $e_1=(1, 0, \cdots, 0), \cdots, e_n=(0, \cdots, 0, 1)$ 是 $\mathfrak{R}_n$ 的酉基底.

练习三十五

1. 证明正定厄米特齐式 $\sum a_{ij}\bar{x}_i x_j$ 满足不等式：

$$(\sum a_{ij}\bar{x}_i y_j)\cdot(\sum a_{ij} x_j \bar{y}_i)\leq(\sum a_{ij}\bar{x}_i x_j)\cdot(\sum a_{ij}\bar{y}_i y_j).$$

2. 取酉空间 $\mathfrak{R}$ 之一组基底 $e_1, \cdots, e_n$．如果 $\mathfrak{R}$ 的任两个向量 $x=\sum_1^n \xi_i e_i$, $y=\sum_1^n \eta_i e_i$ 之数积 $(x, y)=\sum_1^n \xi_i\bar{\eta}_i$，试证：$e_1, \cdots, e_n$ 为 $\mathfrak{R}$ 的酉基底.

第2节　酉变换、酉矩阵、主轴问题

使酉空间 $\mathfrak{R}$ 之任意向量 $\boldsymbol{x}$ 的长保持不变的线性变换 $\mathscr{U}$ 叫酉变换，即 $(\boldsymbol{x}\mathscr{U},\ \boldsymbol{x}\mathscr{U})=(\boldsymbol{x},\ \boldsymbol{x})$.

定理 1　酉空间 $\mathfrak{R}$ 的线性变换 $\mathscr{U}$ 为酉变换的充要条件是对于任两个向量 $\boldsymbol{x}$, $\boldsymbol{y}$, 有 $(\boldsymbol{x}\mathscr{U},\ \boldsymbol{y}\mathscr{U})=(\boldsymbol{x},\ \boldsymbol{y})$.

条件的充分性易证，因为这时令 $\boldsymbol{y}=\boldsymbol{x}$ 就得出 $(\boldsymbol{x}\mathscr{U},\ \boldsymbol{x}\mathscr{U})=(\boldsymbol{x},\ \boldsymbol{x})$，故 $\mathscr{U}$ 是酉变换.

反之，若 $\mathscr{U}$ 为酉变换，则有

$$(\boldsymbol{x}+\lambda\boldsymbol{y},\ \boldsymbol{x}+\lambda\boldsymbol{y})=((\boldsymbol{x}+\lambda\boldsymbol{y})\mathscr{U},\ (\boldsymbol{x}+\lambda\boldsymbol{y})\mathscr{U});$$

再将这等式左右两端分别展开，并利用 $(\boldsymbol{x}\mathscr{U},\ \boldsymbol{x}\mathscr{U})=(\boldsymbol{x},\ \boldsymbol{x})$ 及 $(\boldsymbol{y}\mathscr{U},\ \boldsymbol{y}\mathscr{U})=(\boldsymbol{y},\ \boldsymbol{y})$，就得到

$$\lambda\cdot(\boldsymbol{y},\ \boldsymbol{x})+\bar{\lambda}\cdot(\boldsymbol{x},\ \boldsymbol{y})=\lambda\cdot(\boldsymbol{y}\mathscr{U},\ \boldsymbol{x}\mathscr{U})+\bar{\lambda}\cdot(\boldsymbol{x}\mathscr{U},\ \boldsymbol{y}\mathscr{U});$$

于是，再以 $\lambda=1$ 与 $\lambda=i$ 代入上式中，就分别得到

$$(\boldsymbol{y},\ \boldsymbol{x})+(\boldsymbol{x},\ \boldsymbol{y})=(\boldsymbol{y}\mathscr{U},\ \boldsymbol{x}\mathscr{U})+(\boldsymbol{x}\mathscr{U},\ \boldsymbol{y}\mathscr{U}).$$

与
$$(\boldsymbol{y},\ \boldsymbol{x})-(\boldsymbol{x},\ \boldsymbol{y})=(\boldsymbol{y}\mathscr{U},\ \boldsymbol{x}\mathscr{U})-(\boldsymbol{x}\mathscr{U},\ \boldsymbol{y}\mathscr{U}).$$

再由上两个等式相减，即得 $(\boldsymbol{x},\ \boldsymbol{y})=(\boldsymbol{x}\mathscr{U},\ \boldsymbol{y}\mathscr{U})$，就证明了条件的必要性.

利用定理 1，可与第七章第 3 节定理 2 之证明完全类似地，能证明：

定理 2　设 $\boldsymbol{e}_1$, $\boldsymbol{e}_2$, $\cdots$, $\boldsymbol{e}_n$ 是酉空间 $\mathfrak{R}$ 的酉基底，则 $\mathfrak{R}$ 中线性变换 $\mathscr{U}$ 为酉变换的充要条件是 $\boldsymbol{e}_1\mathscr{U}$, $\boldsymbol{e}_2\mathscr{U}$, $\cdots$, $\boldsymbol{e}_n\mathscr{U}$ 为 $\mathfrak{R}$ 之酉基底.

既然酉空间 $\mathfrak{R}$ 之酉基底不止一组，那末当 $\boldsymbol{e}_1$, $\boldsymbol{e}_2$, $\cdots$, $\boldsymbol{e}_n$ 与 $\boldsymbol{f}_1$, $\boldsymbol{f}_2$, $\cdots$, $\boldsymbol{f}_n$ 是 $\mathfrak{R}$ 的两组酉基底时，联系着这两组酉基底的满秩矩阵如令为 $\boldsymbol{Q}$，即

$$(\boldsymbol{f}_1, \boldsymbol{f}_2, \cdots, \boldsymbol{f}_n) = (\boldsymbol{e}_1, \boldsymbol{e}_2, \cdots, \boldsymbol{e}_n)\cdot\boldsymbol{Q},$$

则当令 $\boldsymbol{x}=\sum\xi_i\boldsymbol{e}_i=\sum\eta_i\boldsymbol{f}_i$ 后，就有

$$(\boldsymbol{x}, \boldsymbol{x}) = \sum\xi_i\bar{\xi}_i = (\xi_1, \cdots, \xi_n)\cdot\begin{pmatrix}\bar{\xi}_1\\ \vdots\\ \bar{\xi}_n\end{pmatrix}$$

$$= (\eta_1, \cdots, \eta_n)\boldsymbol{Q}'\overline{\boldsymbol{Q}}\begin{pmatrix}\bar{\eta}_1\\ \vdots\\ \bar{\eta}_n\end{pmatrix}$$

$$= (\eta_1, \cdots, \eta_n)\begin{pmatrix}\bar{\eta}_1\\ \vdots\\ \bar{\eta}_n\end{pmatrix} = \sum\eta_i\bar{\eta}_i,$$

故有 $\boldsymbol{Q}'\overline{\boldsymbol{Q}}=\boldsymbol{E}$，即 $\boldsymbol{Q}\overline{\boldsymbol{Q}}'=\boldsymbol{E}$.

凡是具有性质 $\boldsymbol{Q}\overline{\boldsymbol{Q}}'=\boldsymbol{E}$ 的矩阵 $\boldsymbol{Q}$ 叫做酉矩阵. 容易证明：酉矩阵有下述性质：

性质 1 $\boldsymbol{U}=(u_{ij})$ 为一 n 级复矩阵，则下述几个条件是两两互为等价的：

i) $\boldsymbol{U}$ 为酉矩阵(即 $\boldsymbol{U}\overline{\boldsymbol{U}}'=\boldsymbol{E}$),

ii) $\boldsymbol{U}'\overline{\boldsymbol{U}}=\overline{\boldsymbol{U}}\boldsymbol{U}'=\boldsymbol{E}$,

iii) $u_{i1}\bar{u}_{j1}+u_{i2}\bar{u}_{j2}+\cdots+u_{in}\bar{u}_{jn}=\delta_{ij}$,

iv) $u_{1i}\bar{u}_{1j}+u_{2i}\bar{u}_{2j}+\cdots+u_{ni}\bar{u}_{nj}=\delta_{ij}$.

条件 iii 是说 $\boldsymbol{U}$ 的任意两个行向量是直交的，且每个行向量的长等于 1，因此，$\boldsymbol{U}$ 之 n 个行向量是 n 复数组空间 $\mathfrak{R}_n$ 的酉基底. 同样，条件 iv 又说明了 $\boldsymbol{U}$ 之 n 个列向量也是 $\mathfrak{R}_n$ 的酉基底.

性质 2 酉矩阵之行列式的模(绝对值)为 1.

事实上，从 $\boldsymbol{U}\overline{\boldsymbol{U}}'=\boldsymbol{E}$ 可知 $1=|\boldsymbol{E}|=|\boldsymbol{U}|\cdot|\overline{\boldsymbol{U}}|$，即得性质 2.

性质 3 设 $\boldsymbol{U}$ 为酉矩阵，则 $\boldsymbol{U}'$, $\boldsymbol{U}^*$, $\boldsymbol{U}^{-1}$ 也都是酉矩

阵；又，若 $\boldsymbol{V}$ 为酉矩阵，则 $\boldsymbol{UV}$ 也是酉矩阵。

证明留给读者自行完成。

关于酉矩阵和酉变换的关系，有：

定理3 酉空间 $\mathfrak{R}$ 中线性变换 $\mathscr{U}$ 为酉变换的充要条件是关于酉基底之矩阵为酉矩阵。

证明 设 $\boldsymbol{e}_1, \boldsymbol{e}_2, \cdots, \boldsymbol{e}_n$ 为酉空间 $\mathfrak{R}$ 之酉基底，线性变换 $\mathscr{U}$ 关于这组酉基底之矩阵令为 $\boldsymbol{U}=(u_{ij})$，于是

$$
\begin{aligned}
(\boldsymbol{e}_i\mathscr{U}, \boldsymbol{e}_j\mathscr{U}) &= \left(\sum_{s=1}^{n} u_{is}\boldsymbol{e}_s, \sum_{t=1}^{n} u_{jt}\boldsymbol{e}_t\right) \\
&= \sum_{s,t=1}^{n} u_{is}\bar{u}_{jt}\cdot(\boldsymbol{e}_s, \boldsymbol{e}_t) \\
&= \sum_{s,t=1}^{n} u_{is}\bar{u}_{jt}\delta_{st} = \sum_{t=1}^{n} u_{it}\bar{u}_{jt}. \qquad (2)
\end{aligned}
$$

若 $\mathscr{U}$ 为酉变换，则据定理2知 $\boldsymbol{e}_1\mathscr{U}, \boldsymbol{e}_2\mathscr{U}, \cdots, \boldsymbol{e}_n\mathscr{U}$ 为酉基底，因而(2)之左边等于 δ_{ij}，故由性质1即知 $\boldsymbol{U}$ 为酉矩阵。反之，若 $\boldsymbol{U}$ 是酉矩阵，则(2)之右边等于 δ_{ij}，因而 $(\boldsymbol{e}_i\mathscr{U}, \boldsymbol{e}_j\mathscr{U})=\delta_{ij}$，即 $\boldsymbol{e}_1\mathscr{U}, \cdots, \boldsymbol{e}_n\mathscr{U}$ 为 $\mathfrak{R}$ 之酉基底。证完。

为了诱导酉空间的主轴问题，先证下述两个引理。

引理1 厄米特矩阵的特征值是实数。

证明 若 α 为厄米特矩阵 $\boldsymbol{A}$ 之特征值，令 $\boldsymbol{X}$ 为对应于 α 的 $\boldsymbol{A}$ 之特征向量，则由 $\boldsymbol{AX}=\alpha\boldsymbol{X}$、$\bar{\boldsymbol{A}}'=\boldsymbol{A}$ 得 $\bar{\alpha}\bar{\boldsymbol{X}}'=\bar{\boldsymbol{X}}'\bar{\boldsymbol{A}}'=\bar{\boldsymbol{X}}'\boldsymbol{A}$，故 $\bar{\alpha}\bar{\boldsymbol{X}}'\boldsymbol{X}=\bar{\boldsymbol{X}}'\boldsymbol{AX}=\alpha\bar{\boldsymbol{X}}'\boldsymbol{X}$，即 $(\alpha-\bar{\alpha})\cdot\bar{\boldsymbol{X}}'\boldsymbol{X}=0$，而 $\bar{\boldsymbol{X}}'\boldsymbol{X}$ 为不等于零之正数，故必 $\alpha=\bar{\alpha}$，这就证得了 α 是实数。

引理2 若 n 元复向量 $\boldsymbol{X}_1=\begin{pmatrix}\xi_1\\ \vdots\\ \xi_n\end{pmatrix}$ 的模 $\|\boldsymbol{X}_1\|=\sqrt{\sum\xi_i\bar{\xi}_i}=1$（即 $(\boldsymbol{X}_1, \boldsymbol{X}_1)=1$），则有一酉矩阵 $\boldsymbol{U}$ 以 $\boldsymbol{X}_1$ 为它的第一个列向量。

下面所讲的不仅是证明 $\boldsymbol{U}$ 的存在，而且连 $\boldsymbol{U}$ 的实际求法也获解决。

即先考虑 n 个未知量 $x_1, x_2, \cdots, x_n$ 的线性齐次方程

$$\overline{\boldsymbol{X}}_1'\boldsymbol{X}=\bar{\xi}_1x_1+\bar{\xi}_2x_2+\cdots+\bar{\xi}_nx_n=0,$$

式中 $\boldsymbol{X}=\begin{pmatrix}x_1\\ \vdots\\ x_n\end{pmatrix}$，取其任一个非零解 $\boldsymbol{Y}_2$，再标准化，得 $\boldsymbol{X}_2=\boldsymbol{Y}_2/\sqrt{\overline{\boldsymbol{Y}}_2'\boldsymbol{Y}_2}$，于是 $\boldsymbol{X}_2$ 为与 $\boldsymbol{X}_1$ 直交的单位向量。

再解 n 个未知量 $x_1, x_2, \cdots, x_n$ 的两个线性无关的线性齐次方程组：

$$\overline{\boldsymbol{X}}_1'\boldsymbol{X}=0 \quad 与 \quad \overline{\boldsymbol{X}}_2'\boldsymbol{X}=0;$$

并取它们的任一非零解 $\boldsymbol{Y}_3$，再标准化得 $\boldsymbol{X}_3=\boldsymbol{Y}_3/\sqrt{\overline{\boldsymbol{Y}}_3'\boldsymbol{Y}_3}$，于是 $\boldsymbol{X}_3$ 为与 $\boldsymbol{X}_1$，$\boldsymbol{X}_2$ 都直交的单位向量。

继续这样作下去，假定已求得了 $n-1$ 个两两直交的单位向量 $\boldsymbol{X}_1, \boldsymbol{X}_2, \cdots, \boldsymbol{X}_{n-1}$；若 $\boldsymbol{Y}_n$ 是齐次方程组

$$\overline{\boldsymbol{X}}_1'\boldsymbol{X}=0,\ \overline{\boldsymbol{X}}_2'\boldsymbol{X}=0,\ \cdots,\ \overline{\boldsymbol{X}}_{n-1}'\boldsymbol{X}=0$$

的任一非零解，则其标准化

$$\boldsymbol{X}_n=\boldsymbol{Y}_n/\sqrt{\overline{\boldsymbol{Y}}_n'\boldsymbol{Y}_n}$$

是和 $\boldsymbol{X}_1, \cdots, \boldsymbol{X}_{n-1}$ 都直交的单位向量。这样，矩阵 $\boldsymbol{U}=(\boldsymbol{X}_1, \boldsymbol{X}_2, \cdots, \boldsymbol{X}_n)$ 即为所求的酉矩阵。证完。

有了引理 2，可证下面的：

定理 4 每个厄米特矩阵 $\boldsymbol{H}$（即 $\overline{\boldsymbol{H}}'=\boldsymbol{H}$）总可以用酉矩阵把它变形为对角阵（即存在酉矩阵 $\boldsymbol{U}$，使 $\overline{\boldsymbol{U}}'\boldsymbol{H}\boldsymbol{U}=\boldsymbol{U}^{-1}\boldsymbol{H}\boldsymbol{U}$ 为对角矩阵）。

证明 用归纳法（关于矩阵 $\boldsymbol{H}$ 的级 n）来证明。假定 λ_1 是 $\boldsymbol{H}$ 的一个特征值，对应于 λ_1，选取 $\boldsymbol{H}$ 的一个特征单位向量 $\boldsymbol{X}_1$；由引理 2，知有一个以 $\boldsymbol{X}_1$ 为第一个列向量的酉矩阵 $\boldsymbol{U}_1=$

$(X_1, X_2, \cdots, X_n)$. 因为

$$(U_1^{-1}H)X_1=U_1^{-1}(HX_1)=\lambda_1 U_1^{-1}X_1=\lambda_1\begin{pmatrix}1\\0\\\vdots\\0\end{pmatrix},$$

故
$$U_1^{-1}HU_1=\begin{pmatrix}\lambda_1 & \alpha_2 & \cdots & \alpha_n\\ 0 & & & \\ \vdots & & H_1 & \\ 0 & & & \end{pmatrix},$$

这里 H_1 是 $n-1$ 级的矩阵.

因为 $\overline{U_1^{-1}HU_1}'=\overline{U}_1'\overline{H}'(\overline{U}_1')^{-1}=U_1^{-1}HU_1$, 故 $U_1^{-1}HU_1$ 是厄米特矩阵, 因而 $\alpha_2=\cdots=\alpha_n=0$, 且 $\overline{H}_1'=H_1$. 于是, 再据归纳法的假设, 存在一个 $n-1$ 级酉矩阵 V 使

$$V^{-1}H_1V=\mathrm{diag}\{\lambda_2, \cdots, \lambda_n\}.$$

这样, 作 n 级矩阵 $U_2=\begin{pmatrix}1 & \\ & V\end{pmatrix}$ 后, 则 U_2 为酉矩阵, 因而 $U_1U_2=U$ 也是酉矩阵, 故

$$\begin{aligned}U^{-1}HU&=U_2^{-1}(U_1^{-1}HU_1)U_2\\&=U_2^{-1}\begin{pmatrix}\lambda_1 & \\ & H_1\end{pmatrix}U_2=\begin{pmatrix}1 & \\ & V^{-1}\end{pmatrix}\begin{pmatrix}\lambda_1 & \\ & H_1\end{pmatrix}\begin{pmatrix}1 & \\ & V\end{pmatrix}\\&=\begin{pmatrix}\lambda_1 & \\ & V^{-1}H_1V\end{pmatrix}=\mathrm{diag}\{\lambda_1, \lambda_2, \cdots, \lambda_n\},\end{aligned}$$

这就证明了定理 4.

若将上述酉矩阵 U 关于酉空间 $\mathfrak{R}$ 之一组酉基底所确定之变换记作 $\mathscr{U}$, 则 $\mathscr{U}$ 就是酉变换, 因而由定理 4 即得其几何的叙述:

定理 5 (酉空间的主轴问题) 设 $e_1, e_2, \cdots, e_n$ 是 n 维酉

空间 $\mathfrak{R}$ 的一组酉基底，$\phi(\xi_1, \xi_2, \cdots, \xi_n)=\sum a_{ij}\bar{\xi}_i\xi_j=\overline{X}'HX$ 是 $\mathfrak{R}$ 在基底 $e_1, e_2, \cdots, e_n$ 内的一个厄米特齐式 $(a_{ij}=\bar{a}_{ji})$. $\xi_1, \xi_2, \cdots, \xi_n$ 是 $\mathfrak{R}$ 中向量 x 关于基底 $e_1, e_2, \cdots, e_n$ 的坐标，即 $x=\sum\xi_i e_i$, 那末空间 $\mathfrak{R}$ 必有一酉变换 $\mathscr{U}$ 把 $e_1, e_2, \cdots, e_n$ 变为另一组酉基底 $f_1, f_2, \cdots, f_n$, 使在这新的酉基底内，所给的厄米特齐式 $\sum a_{ij}\bar{\xi}_i\xi_j=\overline{X}'HX$ 变为典型式

$$\sum\lambda_i\bar{\eta}_i\eta_i=\overline{Y}'\cdot\mathrm{diag}\{\lambda_1, \lambda_2, \cdots, \lambda_n\}\cdot Y,\ Y=\begin{pmatrix}\eta_1\\ \vdots\\ \eta_n\end{pmatrix},$$

这里 $\eta_1, \cdots, \eta_n$ 是向量 x 关于基底 $f_1, \cdots, f_n$ 之坐标，即 $x=\sum\xi_i e_i=\sum\eta_i f_i$.

注意：实际计算主轴问题时，因为对应于厄米特矩阵的两个不同的特征值之特征向量是直交的（证法与第七章第4节的引理3一样），所以也可以引用这来简化直交化的手续，一切具体步骤都完全与第七章讨论欧氏空间的主轴问题时所叙述的一样，这里就不再重复了.

此外，关于厄米特齐式与其惯性指数间之关系，以及关于厄米特矩阵之特征值与特征向量之关系，都有与第七章第4节中定理3到定理5同样的性质，在此就不一一地再列举了，读者可自行补出.

练习三十六

1. 证明：酉矩阵的任一行（列）用模为1之任何数遍乘后，结果还是酉矩阵.

2. 证明下列三条件有二个成立时，另一个也一定成立：i) A 为酉矩阵，ii) A 为厄米特矩阵，iii) $A^2=E$.

3. 试求一酉矩阵 U, 使 $U^{-1}AU$ 为对角阵，已知：

i)
$$A=\begin{pmatrix} \frac{1}{3} & -\frac{1}{3\sqrt{2}} & -\frac{i}{\sqrt{6}} \\ -\frac{1}{3\sqrt{2}} & \frac{1}{6} & \frac{i}{2\sqrt{3}} \\ \frac{i}{\sqrt{6}} & -\frac{i}{2\sqrt{3}} & \frac{1}{2} \end{pmatrix},$$

ii)
$$A=\begin{pmatrix} 0 & i & 1 \\ -i & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}.$$

4. 设 A 是厄米特(或实对称)矩阵，证明存在实数 k，使得当 $\mu\geqslant k$ 时，$\mu E+A$ 总是正定的.

5. 证明：反厄米特(反实对称)矩阵 A 的特征值或为零或为纯虚数，对应于两个不同特征值的特征向量是直交的，有酉矩阵 U 使 $U^{-1}AU$ 成对角形，它的最小多项式无重根，且每特征值的重数等于其特征子空间的维数.

6. 设 P_1 与 P_2 分别是 m 级与 n 级矩阵，$P=\begin{pmatrix} P_1 & A \\ O & P_2 \end{pmatrix}$ 为酉矩阵. 证明 $A=O$ 且 P_1 与 P_2 都是酉矩阵.

7. 设 A 与 B 是两个 n 级满秩矩阵，且 $\bar{A}'A=\bar{B}'B$，试证有一酉矩阵 U，使 $A=UB$.

8. 若 S 为实反对称矩阵，证明 $E-S$ 与 $E+S$ 是满秩的，且 $(E-S)(E+S)^{-1}$ 是直交的.

9. 如上题中，S 为反厄米特矩阵，则有何结论?

第3节 酉矩阵及实直交矩阵的对角形

在第七章介绍了实直交矩阵的概念，在前节又介绍了酉矩阵，它们的一些简单性质也已经谈过了，但它们的特征值、特征向量间的一些很重要的关系还没有触及到. 本节将解决这些问题，并说明在几何学上的应用. 因为实直交矩阵是酉矩阵的特例，所以我们讨论酉矩阵，而把关于实直交矩阵的结果作为推论的形式写出来.

首先，设 α 为酉矩阵 U 之特征值，X_α 是对应于 α 的 U

之任一个特征向量. 由于 $UX_\alpha=\alpha X_\alpha$ 及 $\overline{U}'=U^{-1}$ 就得到

$$\overline{X}'_\alpha\overline{U}'=\bar{\alpha}\overline{X}'_\alpha,\quad 即\quad \overline{X}'_\alpha U^{-1}=\bar{\alpha}\overline{X}'_\alpha$$

于是，

$$\overline{X}'_\alpha U^{-1}UX_\alpha=\bar{\alpha}\overline{X}'_\alpha\cdot\alpha X_\alpha,$$

即

$$\overline{X}'_\alpha X_\alpha=\bar{\alpha}\alpha\overline{X}'_\alpha X_\alpha.$$

因 $\overline{X}'_\alpha X_\alpha\neq 0$ (实际上为一正数), 故 $\bar{\alpha}\alpha=1$, 即酉矩阵 U 的任何特征值的模(绝对值)等于 1.

其次, 若 α 与 β 是酉矩阵 U 之任两个互异的特征值, 而令 X_α 与 X_β 是分别对应于 α 及 β 的 U 之特征向量, 则从 $UX_\alpha=\alpha X_\alpha$, $UX_\beta=\beta X_\beta$ 就得到

$$\overline{X}'_\alpha\overline{U}'=\bar{\alpha}\overline{X}'_\alpha,\quad \overline{X}'_\alpha\overline{U}'UX_\beta=\bar{\alpha}\beta\overline{X}'_\alpha X_\beta,$$

$$(\bar{\alpha}\beta-1)\cdot\overline{X}'_\alpha X_\beta=0;$$

但 $(\bar{\alpha}\beta-1)\alpha=\bar{\alpha}\alpha\beta-\alpha=\beta-\alpha\neq 0$, 故 $\bar{\alpha}\beta-1\neq 0$, 因而必有 $\overline{X}'_\alpha X_\beta=0$, 说明了对应于酉矩阵之不同特征值的特征向量是直交的. 总上所述, 有:

定理 1 酉矩阵的特征值之模(绝对值)等于 1, 对应于不同特征值之特征向量是直交的.

对实直交矩阵, 又有

推论 实直交矩阵之特征值的模等于 1, 对应于两个不同特征值的特征向量是直交的.

利用定理 1, 运用与证明第 2 节中定理 4 的方法, 可以证明:

定理 2 设 U 是酉矩阵, 则必有酉矩阵 V 使 $V^{-1}UV$ 为对角矩阵.

证明 取 U 之任一个特征值 α_1, 则由定理 1 知 $\alpha_1\bar{\alpha}_1=1$. 再选取 U 之对应于 α_1 的特征单位向量 X_1. 于是, 据前节引理 2 知, 有一个以 X_1 为第一个列向量的酉矩阵 $V_1=(X_1, X_2, \cdots, X_n)$. 因之,

$$V_1^{-1}UV_1=\begin{pmatrix}\alpha_1 & s_2 & \cdots & s_n\\ 0 & & & \\ \vdots & & A & \\ 0 & & & \end{pmatrix}.$$

但 $V_1^{-1}UV_1$ 为酉矩阵，故其每行向量的模为 1，因而

$$\alpha_1\bar{\alpha}_1+s_2\bar{s}_2+\cdots+s_n\bar{s}_n=1,$$

于是，由于 $\alpha_1\bar{\alpha}_1=1$，就必致 $s_2=\cdots=s_n=0$，故

$$V_1^{-1}UV_1=\begin{pmatrix}\alpha_1 & 0 & \cdots & 0\\ 0 & & & \\ \vdots & & A & \\ 0 & & & \end{pmatrix},$$

因而 A 是一个 $n-1$ 级的酉矩阵．再利用归纳法，则知有一个 $n-1$ 级酉矩阵 W 使 $W^{-1}AW=\mathrm{diag}\{\alpha_2, \cdots, \alpha_n\}$，因之 $V_2=\begin{pmatrix}1 & \\ & W\end{pmatrix}$ 为 n 级酉矩阵，且有（命 $V=V_1V_2$）

$$V^{-1}UV=\begin{pmatrix}1 & \\ & W^{-1}\end{pmatrix}\begin{pmatrix}\alpha_1 & \\ & A\end{pmatrix}\begin{pmatrix}1 & \\ & W\end{pmatrix}$$
$$=\mathrm{diag}\{\alpha_1, \alpha_2, \cdots, \alpha_n\}.$$

定理 2 证完．

推论　设 A 是实直交矩阵，则必有酉矩阵 V 使 $V^{-1}AV$ 为对角矩阵．当 A 之特征值全是实数（即为 ±1 时），V 可取为实直交矩阵．

既知酉矩阵（实直交矩阵）能与对角矩阵相似，故再利用第五章中已知的结果，又得到：

定理 3　酉矩阵（或实直交矩阵）的最小多项式无重根，因之它的特征矩阵之初等因子都是一次式，又它的每个特征值之重数等于其相应特征子空间的维数．

下面着重来研究实直交矩阵，借以说明它在几何学中的应用。

我们知道：若实直交矩阵 $\boldsymbol{A}$ 的特征值全为实数(即 ± 1)时，由定理 2 之推论得知有实直交矩阵 $\boldsymbol{V}$ 使 $\boldsymbol{V}^{-1}\boldsymbol{A}\boldsymbol{V}$ 为对角矩阵。但是，若 $\boldsymbol{A}$ 之特征值不全为 ± 1 时，这样的实直交矩阵 $\boldsymbol{V}$ 就一定不存在，因为对于复特征值的特征向量不可能是实向量，现在的问题是能否适当地挑选一个实直交矩阵 $\boldsymbol{V}$ 使得 $\boldsymbol{V}^{-1}\boldsymbol{A}\boldsymbol{V}$ 成为某种简单的形状呢？

设 α_1 是 $\boldsymbol{A}$ 之一复数特征值，因为 $|\lambda\boldsymbol{E}-\boldsymbol{A}|=0$ 之系数全都是实数，故 α_1 的共轭复数 $\bar{\alpha}_1$ 也是 $\boldsymbol{A}$ 的特征值，并且 α_1 与 $\bar{\alpha}_1$ 的重复度相同。再若对于 α_1 之特征向量令为 $\boldsymbol{X}_1$，即 $\boldsymbol{A}\boldsymbol{X}_1=\alpha_1\boldsymbol{X}_1$，则 $\boldsymbol{A}\overline{\boldsymbol{X}}_1=\bar{\alpha}_1\overline{\boldsymbol{X}}_1$，说明了 $\boldsymbol{X}_1$ 的共轭向量 $\overline{\boldsymbol{X}}_1$ 是对应于 $\bar{\alpha}_1$ 的特征向量。当 $\|\boldsymbol{X}_1\|=1$ 时，$\|\overline{\boldsymbol{X}}_1\|=1$，也就是说若 $\boldsymbol{X}_1$ 是单位向量，则 $\overline{\boldsymbol{X}}_1$ 亦然。于是可以这样来选择酉矩阵 $\boldsymbol{V}=(\boldsymbol{X}_1, \boldsymbol{X}_2, \boldsymbol{X}_3, \cdots, \boldsymbol{X}_n)$，使 $\boldsymbol{X}_2=\overline{\boldsymbol{X}}_1$ 而有

$$\boldsymbol{V}^{-1}\boldsymbol{A}\boldsymbol{V}=\begin{pmatrix}\alpha_1 & & \\ & \bar{\alpha}_1 & \\ & & \ddots\end{pmatrix}, \quad \text{即} \quad \boldsymbol{A}\boldsymbol{V}=\boldsymbol{V}\begin{pmatrix}\alpha_1 & & \\ & \bar{\alpha}_1 & \\ & & \ddots\end{pmatrix}.$$

因而 $\boldsymbol{A}\boldsymbol{X}_1=\alpha_1\boldsymbol{X}_1$，$\boldsymbol{A}\overline{\boldsymbol{X}}_1=\bar{\alpha}_1\overline{\boldsymbol{X}}_1$。如令 $\alpha_1=\cos\theta_1-i\sin\theta_1$，则 $\bar{\alpha}_1=\cos\theta_1+i\sin\theta_1$，因之

$$\boldsymbol{Y}_1=\frac{\boldsymbol{X}_1+\overline{\boldsymbol{X}}_1}{\sqrt{2}}, \quad \boldsymbol{Y}_2=\frac{\boldsymbol{X}_1-\overline{\boldsymbol{X}}_1}{\sqrt{2}\,i}$$

都是实向量。由定理 1 之推论，$\boldsymbol{V}$ 之前两个列向量 $\boldsymbol{X}_1$ 与 $\overline{\boldsymbol{X}}_1$ 直交，即 $\overline{\boldsymbol{X}}_1'\overline{\boldsymbol{X}}_1=0$，故易证 $\boldsymbol{Y}_1'\boldsymbol{Y}_2=0$，$\|\boldsymbol{Y}_1\|=\|\boldsymbol{Y}_2\|=1$，且 $\boldsymbol{Y}_1$ 与 $\boldsymbol{Y}_2$ 也都和 $\boldsymbol{X}_3, \cdots, \boldsymbol{X}_n$ 直交。这说明了 $\boldsymbol{V}_1=(\boldsymbol{Y}_1, \boldsymbol{Y}_2, \boldsymbol{X}_3, \cdots, \boldsymbol{X}_n)$ 为酉矩阵，其中 $\boldsymbol{Y}_1$ 与 $\boldsymbol{Y}_2$ 是实的。

再因

$$A\boldsymbol{Y}_1=\frac{1}{\sqrt{2}}(A\boldsymbol{X}_1+A\overline{\boldsymbol{X}}_1)=\frac{1}{\sqrt{2}}(\alpha_1\boldsymbol{X}_1+\bar{\alpha}_1\overline{\boldsymbol{X}}_1)$$

$$=\frac{1}{\sqrt{2}}[\cos\theta_1\cdot(\boldsymbol{X}_1+\overline{\boldsymbol{X}}_1)-i\sin\theta_1\cdot(\boldsymbol{X}_1-\overline{\boldsymbol{X}}_1)]$$

$$=\frac{1}{\sqrt{2}}(\sqrt{2}\cos\theta_1\cdot\boldsymbol{Y}_1+\sqrt{2}\sin\theta_1\cdot\boldsymbol{Y}_2)$$

$$=\cos\theta_1\cdot\boldsymbol{Y}_1+\sin\theta_1\cdot\boldsymbol{Y}_2,$$

同理又有 $A\boldsymbol{Y}_2=-\sin\theta_1\cdot\boldsymbol{Y}_1+\cos\theta_1\cdot\boldsymbol{Y}_2$，故

$$A\boldsymbol{V}_1=(A\boldsymbol{Y}_1,\ A\boldsymbol{Y}_2,\ A\boldsymbol{X}_3,\ \cdots,\ A\boldsymbol{X}_n)$$

$$=(\boldsymbol{Y}_1,\ \boldsymbol{Y}_2,\ \boldsymbol{X}_3,\ \cdots,\ \boldsymbol{X}_n)\cdot\begin{pmatrix}\boxed{\begin{matrix}\cos\theta_1 & -\sin\theta_1\\ \sin\theta_1 & \cos\theta_1\end{matrix}} & 0\\ 0 & \ddots\end{pmatrix}$$

$$=\boldsymbol{V}_1\cdot\begin{pmatrix}\boxed{\begin{matrix}\cos\theta_1 & -\sin\theta_1\\ \sin\theta_1 & \cos\theta_1\end{matrix}} & 0\\ 0 & \ddots\end{pmatrix},$$

因之，
$$\boldsymbol{V}_1^{-1}A\boldsymbol{V}_1=\begin{pmatrix}\boxed{\begin{matrix}\cos\theta_1 & -\sin\theta_1\\ \sin\theta_1 & \cos\theta_1\end{matrix}} & 0\\ 0 & \ddots\end{pmatrix}.$$

同样，若 $A\boldsymbol{X}_3=\beta\boldsymbol{X}_3$ 之 β 为复数，则 $\bar{\beta}$ 亦必为 A 之复特征值，且有 $A\overline{\boldsymbol{X}}_3=\bar{\beta}\overline{\boldsymbol{X}}_3$，故可取 $\overline{\boldsymbol{X}}_3=\boldsymbol{X}_4$，于是 $\boldsymbol{Y}_3=\dfrac{\boldsymbol{X}_3+\overline{\boldsymbol{X}}_3}{\sqrt{2}}$ 及 $\boldsymbol{Y}_4=\dfrac{\boldsymbol{X}_3-\overline{\boldsymbol{X}}_3}{\sqrt{2}\,i}$ 都是实的，且有 $\|\boldsymbol{Y}_3\|=\|\boldsymbol{Y}_4\|=1$，及 $\boldsymbol{Y}_3'\boldsymbol{Y}_4=0$. 并易知 $\boldsymbol{Y}_1,\ \boldsymbol{Y}_2,\ \boldsymbol{Y}_3,\ \boldsymbol{Y}_4,\ \boldsymbol{X}_5,\ \cdots,\ \boldsymbol{X}_n$ 为两两直交. 于是，$\boldsymbol{V}_2=(\boldsymbol{Y}_1,\ \boldsymbol{Y}_2,\ \boldsymbol{Y}_3,\ \boldsymbol{Y}_4,\ \boldsymbol{X}_5,\ \cdots,\ \boldsymbol{X}_n)$ 为酉矩阵，其前四个列向量是实的，并有

$$\boldsymbol{A}\boldsymbol{Y}_3=\cos\theta_2\cdot\boldsymbol{Y}_3+\sin\theta_2\cdot\boldsymbol{Y}_4,$$
$$\boldsymbol{A}\boldsymbol{Y}_4=-\sin\theta_2\cdot\boldsymbol{Y}_3+\cos\theta_2\cdot\boldsymbol{Y}_4,$$
但 $\beta=\cos\theta_2-i\sin\theta_2$. 因而

$$\boldsymbol{V}_2^{-1}\boldsymbol{A}\boldsymbol{V}_2=\begin{pmatrix}\boxed{\begin{matrix}\cos\theta_1 & -\sin\theta_1\\ \sin\theta_1 & \cos\theta_1\end{matrix}} & & \\ & \boxed{\begin{matrix}\cos\theta_2 & -\sin\theta_2\\ \sin\theta_2 & \cos\theta_2\end{matrix}} & \\ & & \ddots\end{pmatrix}.$$

继续这样作下去，一直到把 $\boldsymbol{A}$ 之全部 $2k$ 个复特征值用完，剩下来的只是 $\boldsymbol{A}$ 的实特征值 ±1（假如有的话）. 于是得到酉矩阵

$$\boldsymbol{W}=(\boldsymbol{Y}_1,\ \cdots,\ \boldsymbol{Y}_{2k},\ \boldsymbol{X}_{2k+1},\ \cdots,\ \boldsymbol{X}_n),$$

其中前 $2k(\leqslant n)$ 个列向量 $\boldsymbol{Y}_1,\ \cdots,\ \boldsymbol{Y}_{2k}$ 都是实的. 又因 $\boldsymbol{X}_{2k+1},\ \cdots,\ \boldsymbol{X}_n$ 是对应于 $\boldsymbol{A}$ 之实特征值的特征向量，所以也都可以选取为实的. 这样，$\boldsymbol{W}$ 就是实直交矩阵了，这时

$$\boldsymbol{W}^{-1}\boldsymbol{A}\boldsymbol{W}=\begin{pmatrix}\boxed{\begin{matrix}\cos\theta_1 & -\sin\theta_1\\ \sin\theta_1 & \cos\theta_1\end{matrix}} & & & & & & & \\ & \ddots & & & & & & \\ & & \boxed{\begin{matrix}\cos\theta_k & -\sin\theta_k\\ \sin\theta_k & \cos\theta_k\end{matrix}} & & & & & \\ & & & 1 & & & & \\ & & & & \ddots & & & \\ & & & & & 1 & & \\ & & & & & & -1 & \\ & & & & & & & \ddots \\ & & & & & & & & -1\end{pmatrix}$$

于是，又证得了：

定理 4 设 $\boldsymbol{A}$ 为实直交矩阵，则有另一实直交矩阵 $\boldsymbol{T}$，使

$$\boldsymbol{T}^{-1}\boldsymbol{A}\boldsymbol{T}=\begin{pmatrix}1&&&&&&&&\\&\ddots&&&&&&&\\&&1&&&&&&\\&&&-1&&&&&\\&&&&\ddots&&&&\\&&&&&-1&&&\\&&&&&&\boxed{\begin{matrix}\cos\theta_1&-\sin\theta_1\\\sin\theta_1&\cos\theta_1\end{matrix}}&&\\&&&&&&&\ddots&\\&&&&&&&&\boxed{\begin{matrix}\cos\theta_k&-\sin\theta_k\\\sin\theta_k&\cos\theta_k\end{matrix}}\end{pmatrix}.$$

（这又叫做实直交矩阵的实标准形）．

注意：定理 4 中的 $\boldsymbol{T}$ 还可以选取为正常的，即行列式 $|\boldsymbol{T}|=1$．事实上，若 $|\boldsymbol{T}|=-1$，则当 $\boldsymbol{A}$ 有复特征值时，那末在其实标准形中至少有一个象

$$\boxed{\begin{matrix}\cos\theta_1&-\sin\theta_1\\\sin\theta_1&\cos\theta_1\end{matrix}}$$

形状的方块，它在实标准形中所占的两行与两列为同序的；如它是位置在第 k 和第 $k+1$ 行以及在第 k 和第 $k+1$ 列，那末将这两行和两列都互换，其结果等于

$$\boldsymbol{E}_{k,k+1}^{-1}\boldsymbol{T}^{-1}\boldsymbol{A}\boldsymbol{T}\boldsymbol{E}_{k,k+1},$$

但 $\boldsymbol{E}_{k,k+1}$ 是由单位矩阵 $\boldsymbol{E}$ 之第 k，$k+1$ 两列互换所得的初等矩阵，而 $\boldsymbol{T}\boldsymbol{E}_{k,k+1}=\boldsymbol{T}_1$ 也是实直交的且为正常的，这时有

$$
T_1^{-1}AT_1=\begin{pmatrix}1 &&&&&&\\ &\ddots&&&&&\\ &&1&&&&\\ &&&-1&&&\\ &&&&\ddots&&\\ &&&&&-1&\\ &&&&&&\begin{bmatrix}\cos\theta_1' & -\sin\theta_1'\\ \sin\theta_1' & \cos\theta_1'\end{bmatrix}\\ &&&&&&&\ddots\end{pmatrix},
$$

它与定理 4 有相同的实标准形的形状(式中 $\theta_1'=-\theta_1$).

但当 $\boldsymbol{A}$ 只有实特征值 ±1 时, 那只须把 $\boldsymbol{T}$ 之任两个列向量互换, 所得的就是正常实直交矩阵, 这时定理 4 中的标准形最多也不过是有一个 $+1$ 与 -1 之位置互相掉换而已, 还是具有原来的形状. 因此, 定理 4 中的 $\boldsymbol{T}$ 是可以选取为正常的.

下面再讨论实直交矩阵所表示的几何意义.

先讨论二维空间 $\Re_2$. 设 $\boldsymbol{A}=\begin{pmatrix}a_1 & b_1\\ a_2 & b_2\end{pmatrix}$ 是实直交矩阵. 于是

$$
a_1^2+b_1^2=1=a_2^2+b_2^2,\quad a_1a_2+b_1b_2=0.
$$

从 $a_1^2+b_1^2=1$ 可知: 必有(实)角 θ_1, 使

$$
a_1=\cos\theta_1,\quad 因而\quad b_1=\pm\sin\theta_1;
$$

但若 $b_1=\sin\theta_1$, 则令 $\theta'=-\theta_1$ 时就有 $a_1=\cos\theta'$, $b_1=-\sin\theta'$, 总之, 从 $a_1^2+b_1^2=1$, 恒可选取 θ, 使得

$$
a_1=\cos\theta,\quad b_1=-\sin\theta. \tag{3}
$$

于是, 再利用 $a_1^2+a_2^2=1$, 又知 $a_2=\pm\sin\theta$; 故再由 $a_1a_2+b_1b_2=0$ 可知 $b_2=\pm\cos\theta$(当 $\sin\theta\neq0$ 时).

这说明了：当 $\sin\theta\neq 0$ 时，由(3)式又必有

$$a_2=\sin\theta,\quad b_2=\cos\theta$$

或

$$a_2=-\sin\theta,\ b_2=-\cos\theta,$$

二者必有一；也就是说，$\boldsymbol{A}$ 只有下述二种类型：

$$\begin{pmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{pmatrix} \text{ 与 } \begin{pmatrix}\cos\theta & -\sin\theta\\ -\sin\theta & -\cos\theta\end{pmatrix}, \tag{4}$$

其中 θ 为参数．至于 $\sin\theta=0$ 时，$b_1=a_2=0$，因之有 $a_1=\pm1$，且 b_2 也只能为 $+1$ 或 -1，于是

$$\boldsymbol{A}=\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix},\ \text{或}\ \begin{pmatrix}-1 & 0\\ 0 & -1\end{pmatrix},$$

$$\text{或}\ \begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix},\ \text{或}\ \begin{pmatrix}-1 & 0\\ 0 & 1\end{pmatrix};$$

在前二种情形取 (4) 之头一个矩阵之 $\theta=0°$ 或 $180°$，在后二种情形取(4)之后一个矩阵之 $\theta=0°$ 或 $180°$，均可适合．

总之，二级实直交矩阵 $\boldsymbol{A}$ 恒可表写为(4)形．

今将向量 $(x,\ y)$ 看作是平面上的一点 P（即将向量起点恒放在一点，向量$(x,\ y)$可视为点 P 之坐标）．于是

$$\begin{pmatrix}x'\\ y'\end{pmatrix}=\begin{pmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{pmatrix}\begin{pmatrix}x\\ y\end{pmatrix}$$

看作是平面上绕原点旋转 θ 角把点 $P(x,\ y)$变为点 $P'(x',\ y')$ 的变换（参看第四章第 1 节）．这时 $\boldsymbol{A}=\begin{pmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{pmatrix}$ 之行列式 $|\boldsymbol{A}|=1$，所以 $\boldsymbol{A}$ 是正常的．但变换

$$\begin{pmatrix}x'\\ y'\end{pmatrix}=\begin{pmatrix}\cos\theta & -\sin\theta\\ -\sin\theta & -\cos\theta\end{pmatrix}\begin{pmatrix}x\\ y\end{pmatrix}$$

显然是下列两个变换

$$\begin{pmatrix} X \\ Y \end{pmatrix}=\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}$$

与
$$\begin{pmatrix} x' \\ y' \end{pmatrix}=\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\begin{pmatrix} X \\ Y \end{pmatrix}$$

的积，前者表示旋转，后者表示关于 x-轴的反射，故变换

$$\begin{pmatrix} x' \\ y' \end{pmatrix}=\begin{pmatrix} \cos\theta & -\sin\theta \\ -\sin\theta & -\cos\theta \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}.$$

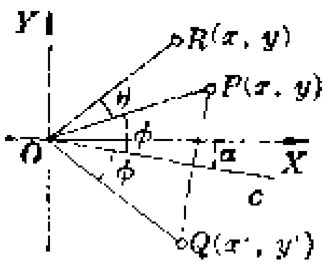
是平面上以直线 $x\cdot\sin\frac{\theta}{2}+y\cdot\cos\frac{\theta}{2}=0$ 为对称轴的反射变换[†]. 这时，$\boldsymbol{A}=\begin{pmatrix} \cos\theta & -\sin\theta \\ -\sin\theta & -\cos\theta \end{pmatrix}$ 为非正常的.

总之，二级实直交矩阵的几何意义，或为绕平面上原点的旋转，或为以通过原点的一直线为对称轴的反射，由它为正常或非正常的来决定.

再讨论三维空间 $\mathfrak{R}_3$. 设 $\boldsymbol{A}$ 是一个三级实直交矩阵. 据定理 4 得知有正常的实直交矩阵 $\boldsymbol{T}$, 使 $\boldsymbol{T}^{-1}\boldsymbol{A}\boldsymbol{T}$ 之形状不外乎下列的六种：

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}=\boldsymbol{E},\quad \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}=\boldsymbol{E}_3(-1),$$

[†] 如右图，连结 PQ，过原点 O 作 PQ 的垂线 OC，即 OC 是 PQ 的中垂线（$\because OP=OR=OQ$），因之 $\phi=\angle QOC=\angle COP$，又 $\phi+\alpha=\angle QOx=\angle xOR=\angle xOP+\theta=\phi-\alpha+\theta$，故 $2\alpha=\theta$，$\alpha=\frac{1}{2}\theta$，即直线 OC 不因 P 点的位置有所变化，它的方程显然是 $x\sin\frac{1}{2}\theta+y\cos\frac{1}{2}\theta=0$。



$$\begin{pmatrix}1 & 0 & 0\\ 0 & -1 & 0\\ 0 & 0 & -1\end{pmatrix}=\boldsymbol{E}_2(-1)\cdot\boldsymbol{E}_3(-1),$$

$$\begin{pmatrix}-1 & 0 & 0\\ 0 & -1 & 0\\ 0 & 0 & -1\end{pmatrix}=-\boldsymbol{E},\quad \begin{pmatrix}1 & 0 & 0\\ 0 & \cos\theta & -\sin\theta\\ 0 & \sin\theta & \cos\theta\end{pmatrix},$$

$$\begin{pmatrix}-1 & 0 & 0\\ 0 & \cos\theta & -\sin\theta\\ 0 & \sin\theta & \cos\theta\end{pmatrix},$$

与此相应的变换分别是

① $\begin{cases}x'=x,\\ y'=y,\\ z'=z;\end{cases}$ ② $\begin{cases}x'=x,\\ y'=y,\\ z'=-z;\end{cases}$

③ $\begin{cases}x'=x,\\ y'=-y,\\ z'=-z;\end{cases}$ ④ $\begin{cases}x'=-x,\\ y'=-y,\\ z'=-z;\end{cases}$

⑤ $\begin{cases}x'=x,\\ y'=y\cos\theta-z\sin\theta,\\ z'=y\sin\theta+z\cos\theta;\end{cases}$ ⑥ $\begin{cases}x'=-x\\ y'=\quad y\cos\theta-z\sin\theta,\\ z'=\quad y\sin\theta+z\cos\theta.\end{cases}$

因为 $\boldsymbol{T}$ 是正常实直交矩阵，故可把 $\boldsymbol{T}$ 看作是将 $\mathfrak{R}_3$ 之原点不动仅将坐标轴(例如，右手系)作适当的旋转而得的新坐标系(仍为右手系)之一变换. 于是，上述的六种点变换都是指关于新坐标系来说的；然矩阵 $\boldsymbol{A}$ 所表示的几何意义，又与坐标轴系之选择无关，这就说明了三维空间 $\mathfrak{R}_3$ 中实直交矩阵所表示的线性变换只限于上述的①～⑥.

显然，①～⑥之几何意义是：

① 恒等变换——空间的每点都不动；

② 以某平面为镜面（即以新坐标面 xOy 为镜面）的反射；

③ 以某直线（即新 x-轴）为轴的反射，即绕这直线旋转 $180°$ 的旋转变换；

④ 以某点（即原点）为对称中心的反射；

⑤ 以某直线（即新 x-轴）为旋转轴而旋转 θ 角的旋转变换；

⑥ 以某直线（即新 x-轴）为轴旋转 θ 角的旋转变换与以垂直这直线的平面（即新坐标面 yOz）为镜面的反射变换之乘积。

前四种可以认为是后二种的特例：因为在⑤中令 $\theta=0$ 即得①，而令 $\theta=\pi$ 即得③；在⑥中令 $\theta=0$ 即得②（不过是将第一、三两个坐标互换了），而令 $\theta=\pi$ 即得④。

所以，三级实直交矩阵 $\boldsymbol{A}$ 之几何意义或者是以一直线为轴的旋转，或者是以一直线为轴之旋转后再伴随一个垂直于此直线之平面为镜面的反射，全由 $\boldsymbol{A}$ 是正常的或非正常的来决定。

练习三十七

1. 设 $\boldsymbol{A}$ 是一个正常三级实直交矩阵。试证：有适合条件 $-1\leqslant t\leqslant 3$ 之数 t，使

$$\boldsymbol{A}^3-t\boldsymbol{A}^2+t\boldsymbol{A}-\boldsymbol{E}=\boldsymbol{0}.$$

2. 设 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 是两个同级的实直交矩阵，并且 $|\boldsymbol{A}|=-|\boldsymbol{B}|$。证明 $\boldsymbol{A}+\boldsymbol{B}$ 是降秩的。

3. 将复矩阵 $\boldsymbol{U}$ 表为 $\boldsymbol{U}=\boldsymbol{P}+i\boldsymbol{Q}$ 之形（$\boldsymbol{P}$、$\boldsymbol{Q}$ 均为实矩阵，$i=\sqrt{-1}$），试证：$\boldsymbol{U}$ 为酉矩阵的充要条件是 $\boldsymbol{P}'\boldsymbol{P}+\boldsymbol{Q}'\boldsymbol{Q}=\boldsymbol{E}$ 及 $\boldsymbol{P}'\boldsymbol{Q}$ 为对称矩阵。

第4节　厄米特齐式耦与(实)二次齐式耦

在第六章里已说过：对于厄米特齐式与（实）二次齐式，可用满秩变换化为典型式，但在某些数学和物理问题中，往往需要把两个二次齐式(即二次齐式耦)同用一个满秩变换化为典型式，这在一般情形是不可能的．例如两个二次齐式

$$\varphi(x_1,\ x_2)=x_1^2,\qquad \psi(x_1,\ x_2)=x_1x_2,$$

若有满秩变换

$$\begin{cases} x_1=c_{11}y_1+c_{12}y_2, \\ x_2=c_{21}y_1+c_{22}y_2 \end{cases} \tag{5}$$

使它们同时化为典型式，则 $\varphi(x_1,x_2)=x_1^2$ 之典型式为 $(c_{11}y_1+c_{12}y_2)^2=\lambda_1y_1^2+\lambda_2y_2^2$ 形，因而系数 c_{11} 与 c_{12} 中至少有一个为零，否则就要出现包含积 y_1y_2 的项 $2c_{11}c_{12}y_1y_2$．不失一般性，可令 $c_{12}=0$，故由(5)是满秩变换，得知 $c_{11}\neq0$，于是这时 $\psi(x_1,x_2)=x_1x_2=c_{11}y_1(c_{21}y_1+c_{22}y_2)=c_{11}c_{21}y_1^2+c_{11}c_{22}y_1y_2$ 由于是典型式，就有 $c_{11}c_{22}=0$，因而必有 $c_{22}=0$，这又与(5)是满秩的假定相矛盾．

然而，在两个二次齐式 $\varphi(x_1,\ \cdots,\ x_n)=\boldsymbol{X}'\boldsymbol{A}\boldsymbol{X}$ 与 $\psi(x_1,\ \cdots,\ x_n)=\boldsymbol{X}'\boldsymbol{B}\boldsymbol{X}$ 中，若有一个是正定的，这种要求是可以实现的．下面研究一般的厄米特齐式，至于实二次齐式是它的特例，就留给读者当练习(参看后面的习题三十八第1、2题)．

这节的主要结果是：

定理　设 $\boldsymbol{A}$, $\boldsymbol{B}$ 是两个厄米特矩阵，且 $\boldsymbol{B}$ 是正定的，则存在一个满秩变换 $\boldsymbol{X}=\boldsymbol{P}\boldsymbol{Y}$，使两个厄米特齐式 $\overline{\boldsymbol{X}}'\boldsymbol{A}\boldsymbol{X}$ 与 $\overline{\boldsymbol{X}}'\boldsymbol{B}\boldsymbol{X}$ 同时分别变成典型式

$$\lambda_1\bar{y}_1y_1+\cdots+\lambda_n\bar{y}_ny_n\quad 与\quad \bar{y}_1y_1+\cdots+\bar{y}_ny_n,$$

这里 $\lambda_1, \cdots, \lambda_n$ 是方程 $|\lambda\boldsymbol{B}-\boldsymbol{A}|=0$ 的根.

证明 由于 $\boldsymbol{B}$ 是正定的，故存在满秩矩阵 $\boldsymbol{Q}$，使 $\overline{\boldsymbol{Q}}'\boldsymbol{B}\boldsymbol{Q}=\boldsymbol{E}$；但由 $\boldsymbol{A}_1=\overline{\boldsymbol{Q}}'\boldsymbol{A}\boldsymbol{Q}$ 之厄米特性，又有酉矩阵 $\boldsymbol{U}_1$ 使 $\overline{\boldsymbol{U}}_1'\boldsymbol{A}_1\boldsymbol{U}_1=\boldsymbol{U}_1^{-1}\boldsymbol{A}_1\boldsymbol{U}_1=\mathrm{diag}\{\lambda_1, \lambda_2, \cdots, \lambda_n\}$. 于是，

$$\overline{\boldsymbol{U}}_1'\overline{\boldsymbol{Q}}'\boldsymbol{B}\boldsymbol{Q}\boldsymbol{U}_1=\overline{\boldsymbol{U}}_1'\boldsymbol{U}_1=\boldsymbol{B},$$

$$\overline{\boldsymbol{U}}_1'\overline{\boldsymbol{Q}}'\boldsymbol{A}\boldsymbol{Q}\boldsymbol{U}_1=\mathrm{diag}\{\lambda_1, \lambda_2, \cdots, \lambda_n\}.$$

故若令 $\boldsymbol{P}=\boldsymbol{Q}\boldsymbol{U}_1$，则 $\boldsymbol{P}$ 满秩且有

$$\overline{\boldsymbol{P}}'\boldsymbol{B}\boldsymbol{P}=\boldsymbol{E}, \quad \overline{\boldsymbol{P}}'\boldsymbol{A}\boldsymbol{P}=\mathrm{diag}\{\lambda_1, \lambda_2, \cdots, \lambda_n\},$$

也就是说，厄米特齐式耦 $\overline{\boldsymbol{X}}'\boldsymbol{A}\boldsymbol{X}$ 与 $\overline{\boldsymbol{X}}'\boldsymbol{B}\boldsymbol{X}$ 能由满秩变换 $\boldsymbol{X}=\boldsymbol{P}\boldsymbol{Y}$ 同时分别变成典型式

$$\lambda_1\bar{y}_1y_1+\lambda_2\bar{y}_2y_2+\cdots+\lambda_n\bar{y}_ny_n$$

与

$$\bar{y}_1y_1+\bar{y}_2y_2+\cdots+\bar{y}_ny_n,$$

定理的主要部分证完了. 剩下来的只是要证明 $\lambda_1, \lambda_2, \cdots, \lambda_n$ 是方程 $|\lambda\boldsymbol{B}-\boldsymbol{A}|=0$ 的根，因为由 $\overline{\boldsymbol{P}}'\boldsymbol{B}\boldsymbol{P}=\boldsymbol{E}$ 及 $\overline{\boldsymbol{P}}'\boldsymbol{A}\boldsymbol{P}=\mathrm{diag}\{\lambda_1, \lambda_2, \cdots, \lambda_n\}$ 即得

$$\begin{aligned}\overline{\boldsymbol{P}}'(\lambda\boldsymbol{B}-\boldsymbol{A})\boldsymbol{P}&=\lambda\boldsymbol{E}-\mathrm{diag}\{\lambda_1, \lambda_2, \cdots, \lambda_n\}\\&=\mathrm{diag}\{\lambda-\lambda_1, \lambda-\lambda_2, \cdots, \lambda-\lambda_n\},\end{aligned}$$

故 $|\overline{\boldsymbol{P}}'\cdot(\lambda\boldsymbol{B}-\boldsymbol{A})\cdot\boldsymbol{P}|=(\lambda-\lambda_1)(\lambda-\lambda_2)\cdots(\lambda-\lambda_n)$，因而

$$|\lambda\boldsymbol{B}-\boldsymbol{A}|=c(\lambda-\lambda_1)(\lambda-\lambda_2)\cdots(\lambda-\lambda_n), \quad c\neq 0,$$

证完.

这定理实质上就是酉空间中主轴问题的推广，因为当 $\boldsymbol{B}=\boldsymbol{E}$ 时，证明中的 $\boldsymbol{Q}$ 可令为 $\boldsymbol{E}$，于是 $\boldsymbol{P}=\boldsymbol{Q}\boldsymbol{U}_1=\boldsymbol{U}_1$ 为酉矩阵，这就变成了前面所谈的主轴问题了.

推论 1 设 $\boldsymbol{A}$、$\boldsymbol{B}$ 是两个厄米特矩阵，且 $\boldsymbol{B}$ 是正定的，那末 $|\lambda\boldsymbol{B}-\boldsymbol{A}|=0$ 的根全为实数.

当 $\boldsymbol{B}=\boldsymbol{E}$ 时就得到厄米特矩阵之特征值都是实数，因此这性质是这推论 1 的特例.

注意：推论 1 还可给出它的一个直接的证明．事实上，若 α 为 $|\lambda\boldsymbol{B}-\boldsymbol{A}|=0$ 之任一个根，则由 $(\alpha\boldsymbol{B}-\boldsymbol{A})\boldsymbol{X}=0$ 之一非零解 $\boldsymbol{X}$ 得 $\boldsymbol{AX}=\alpha\boldsymbol{BX}$，故 $\overline{\boldsymbol{X}}'\overline{\boldsymbol{A}}'=\bar{\alpha}\overline{\boldsymbol{X}}'\overline{\boldsymbol{B}}'$，$\overline{\boldsymbol{X}}'\boldsymbol{A}=\bar{\alpha}\overline{\boldsymbol{X}}'\boldsymbol{B}$，$\overline{\boldsymbol{X}}'\boldsymbol{AX}=\bar{\alpha}\overline{\boldsymbol{X}}'\boldsymbol{BX}$，$\alpha\overline{\boldsymbol{X}}'\boldsymbol{BX}=\bar{\alpha}\overline{\boldsymbol{X}}'\boldsymbol{BX}$，即 $(\alpha-\bar{\alpha})\cdot\overline{\boldsymbol{X}}'\boldsymbol{BX}=0$．由 $\boldsymbol{B}$ 之正定性及 $\boldsymbol{X}\neq 0$，得 $\overline{\boldsymbol{X}}'\boldsymbol{BX}>0$，故 $\alpha=\bar{\alpha}$ 为实数．

推论 2 设 $\boldsymbol{A}$、$\boldsymbol{B}$ 是两个厄米特矩阵，且 $\boldsymbol{B}$ 是正定的．如果 α 是 $|\lambda\boldsymbol{B}-\boldsymbol{A}|=0$ 的一个 k 重根，则线性齐次方程组 $(\alpha\boldsymbol{B}-\boldsymbol{A})\boldsymbol{X}=\boldsymbol{0}$ 的解空间是 $\mathfrak{R}_n$ 的 k 维子空间．

这是因为，令 $\boldsymbol{D}=\operatorname{diag}\{\lambda_1, \lambda_2, \cdots, \lambda_n\}$ 时，由于 α 在 $\lambda_1, \lambda_2, \cdots, \lambda_n$ 中要出现 k 次，故 $r_{(\alpha\boldsymbol{E}-\boldsymbol{D})}=n-k$．再据上定理，知道有满秩矩阵 $\boldsymbol{P}$，使 $\overline{\boldsymbol{P}}'\boldsymbol{BP}=\boldsymbol{E}$ 及 $\overline{\boldsymbol{P}}'\boldsymbol{AP}=\boldsymbol{D}$，于是 $r_{(\alpha\boldsymbol{B}-\boldsymbol{A})}=r_{\overline{\boldsymbol{P}}'(\alpha\boldsymbol{B}-\boldsymbol{A})\boldsymbol{P}}=r_{(\alpha\boldsymbol{E}-\boldsymbol{D})}=n-k$，故 $(\alpha\boldsymbol{B}-\boldsymbol{A})\boldsymbol{X}=\boldsymbol{0}$ 之解空间是 k 维的．证完．

练习三十八

1．设 $\boldsymbol{X}'\boldsymbol{AX}$ 与 $\boldsymbol{X}'\boldsymbol{BX}$ 是两个实二次齐式，且 $\boldsymbol{X}'\boldsymbol{BX}$ 是正定的．试证：有满秩变换(实的) $\boldsymbol{X}=\boldsymbol{PY}$，使上述两个实二次齐式分别变为

$$\lambda_1 y_1^2+\lambda_2 y_2^2+\cdots+\lambda_n y_n^2, \quad y_1^2+y_2^2+\cdots+y_n^2,$$

这里的 $\lambda_1, \cdots, \lambda_n$ 是 $|\lambda\boldsymbol{B}-\boldsymbol{A}|=0$ 之根，且为实数．

2．设 $\boldsymbol{A}$、$\boldsymbol{B}$ 是二个同级的正定厄米特矩阵，则 $|\lambda\boldsymbol{B}-\boldsymbol{A}|=0$ 之根全为正实数．

3．凡适合条件 $\overline{\boldsymbol{A}}'\boldsymbol{A}=\boldsymbol{A}\overline{\boldsymbol{A}}'$ 之矩阵 $\boldsymbol{A}$ 叫规范矩阵．试证：对角矩阵、酉矩阵(实直交矩阵)、厄米特矩阵(实对称矩阵)、反厄米特矩阵(实反对称矩阵)都是规范矩阵．又，若 $\boldsymbol{A}$ 是规范矩阵，$\boldsymbol{U}$ 为酉矩阵，则 $\overline{\boldsymbol{U}}'\boldsymbol{AU}$ 也是规范的．

4．$\boldsymbol{A}$ 为规范矩阵的充要条件是存在一个酉矩阵 $\boldsymbol{U}$，使 $\boldsymbol{U}^{-1}\boldsymbol{AU}$ 为对角矩阵．

5．设 $\boldsymbol{A}$ 是规范矩阵，$\lambda_1, \lambda_2, \cdots, \lambda_n$ 是 $\boldsymbol{A}$ 之所有特征值．试证：

i) 若每 λ_i 为实数，则 $\boldsymbol{A}$ 是厄米特的；

ii) 若每 λ_i 为 0 或纯虚数，则 $\boldsymbol{A}$ 是反厄米特的；

iii) 若每 λ_i 的模为 1，则 $\boldsymbol{A}$ 是酉矩阵．

6. 设规范矩阵 $\boldsymbol{A}$ 之 n 个特征值为 $\lambda_1, \lambda_2, \cdots, \lambda_n$，并令 $\mu_1, \cdots, \mu_n$ 为 $\overline{A}'A$ 之特征值. 证明

$$\mu_1+\cdots+\mu_n=|\lambda_1|^2+\cdots+|\lambda_n|^2.$$

7. 试证：三角矩阵为规范的充要条件是它为对角矩阵.

8. 对任何矩阵 $\boldsymbol{A}$(满秩或降秩)，总有一个酉矩阵 $\boldsymbol{U}$，使 $\boldsymbol{U}^{-1}\boldsymbol{A}\boldsymbol{U}$ 为三角矩阵.

9. 将两个二次齐式 $-y^2+2yz-2xy$ 与 $2x^2+2y^2+3z^2-4yz-4zx+2xy$ 同时化为典型式，并求所需的满秩变换.

10. 设 $\boldsymbol{U}$ 为 n 级酉矩阵，$\boldsymbol{A}=\begin{pmatrix} a_1 & & 0 \\ & a_2 & \\ & & \ddots \\ 0 & & a_n \end{pmatrix}$，试证 $\boldsymbol{UA}$ 的特征值 ω 恒满足关系式

$$m\leqslant|\omega|\leqslant M,$$

其中 $M=\max\limits_{1\leqslant r\leqslant n}|a_r|$，$m=\min\limits_{1\leqslant r\leqslant n}|a_r|$.

11. 设 $\boldsymbol{A}$ 为酉矩阵，$\boldsymbol{B}$ 为规范矩阵 (即 $\overline{\boldsymbol{B}}'\boldsymbol{B}=\boldsymbol{B}\overline{\boldsymbol{B}}'$)；而 m 与 M 分别为 $\boldsymbol{B}$ 之特征值中绝对值最小的数与最大的数. 若 $|\boldsymbol{A}-\omega\boldsymbol{B}|=0$，试证 $m\leqslant|\omega|^{-1}\leqslant M$.

第九章 矩阵分析

在以往的各章里，我们都只限于使用四种有理运算（加、减、乘、除），偶而也用过求多项式之根（因而牵涉到开多少次方的运算），但完全没有涉及极限的运算. 本章将研究矩阵的极限运算，先讨论几个常用的矩阵函数 $e^{\boldsymbol{A}}$、$\sin\boldsymbol{A}$、$\cos\boldsymbol{A}$，再利用它们去解线性微分方程组. 本章里的 $\boldsymbol{A}$、$\boldsymbol{B}$ 等都是指 n 级矩阵（数量的），矩阵 $\boldsymbol{A}$, $\boldsymbol{B}$, … 的第 i 行第 j 列处的元素用符号 $(\boldsymbol{A})_{ij}$, $(\boldsymbol{B})_{ij}$, … 来表示，今后都不另说明. 最后，介绍利用矩阵的极限来求线性方程组之近似解，也就是线代数计算中所谓简单迭代法与逐步迭代法收敛的充要条件.

第1节 收敛的矩阵序列

与普通数学分析一样，矩阵的分析理论也是建立在收敛的概念之上的，首先有

定义1 设有矩阵序列 $\{\boldsymbol{A}_m\}$，如果当 $m\to\infty$ 时恒有关系式 $(\boldsymbol{A}_m)_{ij}\to(\boldsymbol{A})_{ij}$, $(i, j=1, \cdots, n)$，我们就叫序列 $\{\boldsymbol{A}_m\}$ 收敛于矩阵 $\boldsymbol{A}$，表以 $\boldsymbol{A}_m\to\boldsymbol{A}$ 或 $\lim\limits_{m\to\infty}\boldsymbol{A}_m=\boldsymbol{A}$，并叫 $\boldsymbol{A}$ 为 $\{\boldsymbol{A}_m\}$ 之极限. 不为收敛的矩阵序列叫做发散的.

换句话说，所谓 $\{\boldsymbol{A}_m\}$ 收敛于 $\boldsymbol{A}$，指的是 $\boldsymbol{A}_m$ 的每个元素当 $m\to\infty$ 时以 $\boldsymbol{A}$ 之对应位置的元素为它的极限.

关于矩阵序列收敛的性质，有许多是和收敛数列类似的. 例如，当 $m\to\infty$ 时，若 $\boldsymbol{A}_m\to\boldsymbol{A}$, $\boldsymbol{B}_m\to\boldsymbol{B}$，则必有

$$\alpha A_m+\beta B_m\to\alpha A+\beta B \tag{1}$$

及

$$A_mB_m\to AB, \tag{2}$$

式中 α、β 为任何的数.

(1)式之正确性可直接由定义得知．需要证明的是(2)式：事实上，由 $A_m\to A$ 及 $B_m\to B$ 易知

$$(A_mB_m)_{ij}=\sum_{t=1}^{n}(A_m)_{it}\cdot(B_m)_{tj}\to\sum_{t=1}^{n}A_{it}B_{tj}=(AB)_{ij}.$$

即 $A_mB_m\to AB$.

由(2)式又立即推得

$$PA_mQ\to PAQ. \tag{3}$$

我们知道，在数学分析中，一个数列是否有极限，是紧密地联系着"无穷小"这个概念的，即：数列 $\{a_m\}$ 有极限 a 的充要条件是数列$\{a_m-a\}$为无穷小量，即 $\lim\limits_{m\to\infty}(a_m-a)=0$. 因之，研究收敛矩阵序列时，也有类似的问题．实际上，由定义易知，矩阵序列 $\{A_m\}\to A$ 的充要条件是$\{B_m\}\to O$，其中 $B_m=A_m-A$. 于是，首先来解决一矩阵序列 $\{A_m\}$ 为无穷小（即 $\lim\limits_{m\to\infty}A_m=O$）的充要条件是什么？

设有矩阵序列 $\{A_m\}$. 首先，若 $\lim\limits_{m\to\infty}A_m=O$，则 $\lim\limits_{m\to\infty}(A_m)_{ij}=0(i=1,\ \cdots,\ n;\ j=1,\ \cdots,\ n)$，故 $\lim\limits_{m\to\infty}\max|(A_m)_{ij}|=0$. 反之，若 $\lim\limits_{m\to\infty}\max\limits_{1\leqslant i,j\leqslant n}|(A_m)_{ij}|=0$，则当然有 $\lim\limits_{m\to\infty}(A_m)_{ij}=0$，即 $\lim\limits_{m\to\infty}A_m=O$. 故 $\{A_m\}\to O$ 之充要条件是 $\max\limits_{1\leqslant i,j\leqslant n}|(A_m)_{ij}|\to 0$.

定义 2 对任意矩阵 $A=(a_{ij})$，定义 $\|A\|=\max\limits_{1\leqslant i,j\leqslant n}|a_{ij}|$ 而叫它为 A 的范数.

于是，总括上述，证得了：

定理 1 矩阵序列 $\{A_m\}$ 为无穷小之充要条件是数列 $\{\|A_m\|\}$ 为无穷小量，即 $\lim\limits_{m\to\infty} A_m = O$ 当且仅当 $\lim\limits_{m\to\infty}\|A_m\|=0$ 时.

关于矩阵的范数，易证：

性质 1° 对 k 个 n 级矩阵 A_1、A_2、…、A_k，恒有

$$\|A_1A_2\cdots A_k\|\leqslant n^{k-1}\|A_1\|\cdot\|A_2\|\cdots\|A_k\|.$$

证明 今采用数学归纳法来证明. $k=1$ 时显然成立. 如假定 $\|A_1A_2\cdots A_{k-1}\|\leqslant n^{k-2}\cdot\|A_1\|\cdot\|A_2\|\cdots\|A_{k-1}\|$，于是，若令 $A_1A_2\cdots A_{k-1}=A=(a_{ij})$，$A_k=B=(b_{ij})$，则 $\|A_1A_2\cdots A_{k-1}A_k\| =\|AB\|=\max\limits_{1\leqslant i,j\leqslant n}|(AB)_{ij}|=\max\limits_{1\leqslant i,j\leqslant n}\left|\sum\limits_{t=1}^{n}a_{it}b_{tj}\right|\leqslant\max\limits_{1\leqslant i,j\leqslant n}\left(\sum\limits_{t=1}^{n}|a_{it}| \cdot|b_{tj}|\right)\leqslant\sum\limits_{t=1}^{n}\max\limits_{1\leqslant i,j\leqslant n}|a_{ij}|\cdot\max\limits_{1\leqslant i,j\leqslant n}|b_{ij}| = n\|A\|\cdot\|B\| = n\|A_1\cdots A_{k-1}\|\cdot\|A_k\|\leqslant n^{k-1}\|A_1\|\|A_2\|\cdots\|A_{k-1}\|\cdot\|A_k\|$，因之性质 1° 获证.

关于数列的极限，常引用由一个数之幂所成的数列，即考虑由 x, x^2, x^3, …, x^m, … 所构成之数列 $\{x^m\}$ 的极限. 我们知道，$\{x^m\}$ 有极限之充要条件是 $|x|<1$，且这时 $\lim\limits_{m\to\infty}x^m=0$. 相应地，在矩阵序列中，首先要考虑的也是一个由矩阵 A 之幂 A, A^2, A^3, …, A^m, … 所构成的矩阵序列 $\{A^m\}$. 为研究的方便，考察 A 之某类标准形，如令

$$P^{-1}AP=J=\begin{pmatrix}\lambda_1 & & *_1\\ & \lambda_2 & \\ 0 & & \ddots \\ & & & \lambda_n\end{pmatrix}$$

为 A 之约旦标准形，因之 λ_1, λ_2, …, λ_n 为 A 之特征值，故

$$P^{-1}A^mP=J^m=\begin{pmatrix}\lambda_1^m & & & *_m\\ & \lambda_2^m & & \\ & & \ddots & \\ 0 & & & \lambda_n^m\end{pmatrix}.$$

于是，据(3)式可知$\{A^m\}$有极限的充要条件是$\{J^m\}$有极限，因之问题转化为求$\{J^m\}$有极限之充要条件.

若$\{J^m\}$有极限，即$\lim\limits_{m\to\infty}J^m$存在，则由定义可知对每个$i$有$\lim\limits_{m\to\infty}\lambda_i^m$必存在，因之条件为$|\lambda_i|<1(i=1, \cdots, n)$，且$\lim\limits_{m\to\infty}\lambda_i^m=0$.

反之，再设$|\lambda_i|<1(i=1, \cdots, n)$. 如同在数学分析中处理这类问题时一样，取适合$|\lambda_i|\leqslant\omega_i<1$之一正数$\omega_i(i=1, \cdots, n)$，并作矩阵

$$T=\begin{pmatrix}\omega_1 & & & *'\\ & \omega_2 & & \\ & & \ddots & \\ 0 & & & \omega_n\end{pmatrix},$$

其目的是为将$\{J^m\}$之研究转化为对$\{T^m\}$之研究，并使后者的研究要方便些；为此，先令$\omega_1, \omega_2, \cdots, \omega_n$两两互异（这很容易办到，因为适合$|\lambda_i|\leqslant\omega_i<1$之$\omega_i$有无限多)，则$T$与对角矩阵相似，即有满秩矩阵$S$使$S^{-1}TS=D=\mathrm{diag}\{\omega_1, \omega_2, \cdots, \omega_n\}$，于是$S^{-1}T^mS=D^m=\mathrm{diag}\{\omega_1^m, \omega_2^m, \cdots, \omega_n^m\}$；而$D^m\to O$($m\to\infty$时）是很显然的，故由(3)得$\{T^m\}\to O$，这说明了不论怎样选$*'$恒有$\{T^m\}\to O$，只要适合$n$个不等式$|\lambda_i|\leqslant\omega_i<1(i=1, \cdots, n)$之$\omega_1, \omega_2, \cdots, \omega_n$两两互异就行了. 于是，剩下来的问题是如何选取$*'$，使得$T$与$J$发生某些必要的关系，便于从$\lim\limits_{m\to\infty}T^m=O$能看出$\lim\limits_{m\to\infty}J^m$的存在. 为此，不妨令

$\boldsymbol{T}$中在 $*'$ 处的各数都是正的且不小于 $\boldsymbol{J}$ 中在 $*_1$ 处的对应位置之数的模．象这样的每个元素都不是负数的矩阵 $\boldsymbol{T}$ 叫做非负矩阵，而 $\boldsymbol{J}$ 之每元素的模都不大于 $\boldsymbol{T}$ 之对应位置的元素，这一事实可简记为 $\boldsymbol{J}\ll\boldsymbol{T}$.

定义 3 设 $\boldsymbol{A}=(a_{ij})$ 是任一个矩阵（其中 a_{ij} 一般表示复数），而 $\boldsymbol{B}=(b_{ij})$ 为一个非负矩阵（即每 $b_{ij}\geqslant 0$）．若 $|a_{ij}|\leqslant b_{ij}(i,\ j=1,\ 2,\ \cdots,\ n)$，就叫 $\boldsymbol{B}$ 优于 $\boldsymbol{A}$（或 $\boldsymbol{A}$ 劣于 $\boldsymbol{B}$），用记号 $\boldsymbol{A}\ll\boldsymbol{B}$ 表示之.

性质 2° 设 $\boldsymbol{A}$ 为一个（复）矩阵，$\boldsymbol{B}$ 为非负矩阵，m 是正整数．那末 $\boldsymbol{B}^m$ 为非负矩阵，且在 $\boldsymbol{A}\ll\boldsymbol{B}$ 时有 $\boldsymbol{A}^m\ll\boldsymbol{B}^m$.

证明 因 $\boldsymbol{B}$ 为非负矩阵，由矩阵乘法的定义，即得 $\boldsymbol{B}^m$ 为非负的．对于第二个论断，也易于理解：因为 $m=1$ 时显然是正确的，再归纳地假定 $\boldsymbol{A}^{m-1}\ll\boldsymbol{B}^{m-1}$，则

$$
\begin{aligned}
|(\boldsymbol{A}^m)_{ij}| &\leqslant \sum_{t=1}^{n} |(\boldsymbol{A}^{m-1})_{it}|\cdot|(\boldsymbol{A})_{tj}| \\
&\leqslant \sum_{t=1}^{n} (\boldsymbol{B}^{m-1})_{it}\cdot(\boldsymbol{B})_{tj}=(\boldsymbol{B}^m)_{ij},
\end{aligned}
$$

即证得了 $\boldsymbol{A}^m\ll\boldsymbol{B}^m$.

性质 3° 设 $\{\boldsymbol{A}_m\}$ 是矩阵序列，$\{\boldsymbol{B}_m\}$ 为非负矩阵的序列，并有 $\boldsymbol{A}_m\ll\boldsymbol{B}_m$ 与 $\boldsymbol{B}_m\to\boldsymbol{O}$．那末 $\boldsymbol{A}_m\to\boldsymbol{O}$.

证明 由 $\boldsymbol{A}_m\ll\boldsymbol{B}_m$ 得 $|(\boldsymbol{A}_m)_{ij}|<(\boldsymbol{B}_m)_{ij}$；由 $\boldsymbol{B}_m\to\boldsymbol{O}$ 又知道：当 $m\to\infty$ 时有 $(\boldsymbol{B}_m)_{ij}\to 0$，因之当 $m\to\infty$ 时有 $|(\boldsymbol{A}_m)_{ij}|\to 0$，随而必有 $(\boldsymbol{A}_m)_{ij}\to 0$，即 $\boldsymbol{A}_m\to\boldsymbol{O}$.

性质 4° 若 $\boldsymbol{A}$ 为复矩阵，且 $\boldsymbol{B}$ 为非负的，而 $\boldsymbol{A}\ll\boldsymbol{B}$，则 $\|\boldsymbol{A}\|\leqslant\|\boldsymbol{B}\|$.

证明 若令 $\boldsymbol{A}=(a_{ij})$，$\boldsymbol{B}=(b_{ij})$，则由 $\boldsymbol{A}\ll\boldsymbol{B}$ 即得 $|a_{ij}|\leqslant b_{ij}(i,\ j=1,\ 2,\ \cdots,\ n)$，于是

$$\|A\| = \max_{1\leqslant i,j\leqslant n} |a_{ij}| \leqslant \max_{1\leqslant i,j\leqslant n} b_{ij} = \|B\|.$$

现在再回到上面的矩阵序列 $\{J^m\}$ 与 $\{T^m\}$ 的讨论. 因为 $J \ll T$, 故由性质 2° 有 $J^m \ll T^m$; 但因 $\lim\limits_{m\to\infty} T^m = O$, 故据性质 3° 则知 $\lim\limits_{m\to\infty} J^m = O$.

总之, $\{J^m\}$有极限之充要条件是 $|\lambda_i|<1$, 且这时 $\lim\limits_{m\to\infty} J^m = 0$, 随而 $\lim\limits_{m\to\infty} A^m = O$. 即证得了:

定理 2 由矩阵 A 之幂 $A, A^2, \cdots, A^m, \cdots$ 构成的矩阵序列 $\{A^m\}$ 有极限(即 $\lim\limits_{m\to\infty} A^m$ 存在)的充要条件是 A 的所有特征值之模都小于 1, 这时必有 $\lim\limits_{m\to\infty} A^m = O$.

附注: 定理 2 的证明中, 证明必要性很简单, 可是证充分性时, 虽然思路是明确的, 但证明过程太长. 于是问: 有直接简单之证法吗? 若仔细回顾上述的证明过程, 容易发现上述充分条件之证明冗长的原因. 因为证必要条件时是先假定了 $\lim\limits_{m\to\infty} J^m$ 存在, 从而保证必有 $|\lambda_i|<1$, 至于在 $*_m$ 处之元素就没有考虑的必要. 反之, 当每 $|\lambda_i|<1$ 时, 虽然 $\lim\limits_{m\to\infty} \lambda_i^{m_i}=0$, 而 $*_m$ 处之元素是什么则无从获知, 故不能立即断言 $\lim\limits_{m\to\infty} J^m$ 之存在, 因之需转化为 $\{T^m\}$ 的研究, 且转化了之后从 $\lim\limits_{m\to\infty} T^m = O$ 就能断言 $\lim\limits_{m\to\infty} J^m = O$, 这说明了不仅 $\lim\limits_{m\to\infty} J^m$ 存在, 且还等于 O, 这样才可肯定 $*_m$ 之各元之极限亦必为 0. 抓住这一个线索, 我们设法计算出 $*_m$ 处之元素, 从而可直接求 $\{J^m\}$ 之极限, 具体方法如下:

关于定理 2 中条件充分性的另一证明 因 J 为 A 之约旦标准形, 故

$$\boldsymbol{J}=\begin{pmatrix}\boldsymbol{J}_1 & & 0\\ & \ddots & \\ 0 & & \boldsymbol{J}_s\end{pmatrix},$$

每 $\boldsymbol{J}_i$ 为 $\boldsymbol{A}$ 之约旦块. 例如设

$$\boldsymbol{J}_1=\begin{pmatrix}\lambda_1 & 1 & 0 & \cdots & 0 & 0\\ 0 & \lambda_1 & 1 & \cdots & 0 & 0\\ 0 & 0 & \lambda_1 & \cdots & 0 & 0\\ \vdots & \vdots & \vdots & & \vdots & \vdots\\ 0 & 0 & 0 & \cdots & \lambda_1 & 1\\ 0 & 0 & 0 & \cdots & 0 & \lambda_1\end{pmatrix}$$

为 k 级的, 由归纳法易证

$$\boldsymbol{J}_1^m=\begin{pmatrix}\lambda_1^m & C_m^1\lambda_1^{m-1} & \cdots & \cdots & \cdots & C_m^{k-1}\lambda_1^{m-k+1}\\ & \lambda_1^m & C_m^1\lambda_1^{m-1} & \cdots & \cdots & C_m^{k-2}\lambda_1^{m-k+2}\\ & & \lambda_1^m & \cdots & \cdots & C_m^{k-3}\lambda_1^{m-k+3}\\ & & & \ddots & & \vdots\\ & & & & \lambda_1^m & C_m^1\lambda_1^{m-1}\\ & & & & & \lambda_1^m\end{pmatrix},$$

式中 C_m^l 为从 m 个事物中取 l 个的组合之总数, 并假定 $C_m^l=0$ (当 $m<l$ 时). 于是, 当 $m\geqslant k$ 时, 则知上述的上三角矩阵 $\boldsymbol{J}_1^m$ 中非零元素均为 λ_1 之某正指数幂的正整数倍, 因之当 $|\lambda_1|<1$ 时有 $\lim\limits_{m\to\infty}\boldsymbol{J}_1^m=\boldsymbol{O}$. 对其它约旦块也有相同的结论. 故当每个 $|\lambda_i|<1$ 时 $(i=1, 2, \cdots, n)$, 则从

$$\boldsymbol{J}^m=\begin{pmatrix}\boldsymbol{J}_1^m & & \\ & \boldsymbol{J}_2^m & \\ & & \ddots & \\ & & & \boldsymbol{J}_s^m\end{pmatrix}$$

即知

$$\lim_{m\to\infty} \boldsymbol{J}^m = \begin{pmatrix} \lim \boldsymbol{J}_1^m & & & \\ & \lim \boldsymbol{J}_2^m & & \\ & & \ddots & \\ & & & \lim \boldsymbol{J}_s^m \end{pmatrix} = \begin{pmatrix} \boldsymbol{O} & & & \\ & \boldsymbol{O} & & \\ & & \ddots & \\ & & & \boldsymbol{O} \end{pmatrix} = \boldsymbol{O},$$

因之条件的充分性获证.

第 2 节 矩阵的幂级数

幂级数的理论,在分析学中占有很重要的地位,现在我们研究矩阵分析的理论时,也特别着重矩阵幂级数这个问题.当然,谈到矩阵的级数,首先应该说明收敛与发散的概念,这和数项级数的定义及其一些基本性质完全类似.

定义 1 给了矩阵级数

$$\sum_{m=1}^{\infty} \boldsymbol{A}_m, \tag{4}$$

令 $\boldsymbol{S}_N = \sum_{m=1}^{N} \boldsymbol{A}_m$. 若矩阵序列$\{\boldsymbol{S}_N\}$收敛而有极限$\boldsymbol{S}$,即 $\lim_{N\to\infty} \boldsymbol{S}_N = \boldsymbol{S}$,就叫级数(4)收敛并有和 $\boldsymbol{S}$,用 $\boldsymbol{S} = \sum_{m=1}^{\infty} \boldsymbol{A}_m$ 来表示.不是收敛的矩阵级数叫做发散的.

显然,$\sum_{m=1}^{\infty} \boldsymbol{A}_m = \boldsymbol{S}$ 的意义指的是

$$\sum_{m=1}^{\infty} (\boldsymbol{A}_m)_{ij} = (\boldsymbol{S})_{ij} \quad (i,\ j=1,\ 2,\ \cdots,\ n). \tag{5}$$

定义 2 如果(5)式左端 n^2 个数项级数都是绝对收敛,就叫矩阵级数(4)是绝对收敛的.

于是,从这定义以及分析学中相应的结果,即得:

性质 1° 若矩阵级数(4)绝对收敛,则它也一定收敛,并且任意调换各项次序所得的新级数还是收敛的,而且和也

不变。

判别(4)是否为绝对收敛，有下面的准则，即

性质 2° 级数 $\sum_{m=1}^{\infty}\boldsymbol{A}_m$ 绝对收敛的充要条件是 $\sum_{m=1}^{\infty}\|\boldsymbol{A}_m\|$ 为收敛级数。

证明 先设 $\sum_{m=1}^{\infty}\boldsymbol{A}_m$ 绝对收敛。于是，存在正数 K，它与 N、i、j 都没有关系，并有

$$\sum_{m=1}^{N}|(\boldsymbol{A}_m)_{ij}|<K \quad (N\geqslant 1;\ i,\ j=1,\ 2,\ \cdots,\ n).$$

因之
$$\sum_{m=1}^{N}\|\boldsymbol{A}_m\|\leqslant\sum_{m=1}^{N}\sum_{i,j=1}^{n}|(\boldsymbol{A}_m)_{ij}|<n^2K,$$

故 $\sum_{m=1}^{\infty}\|\boldsymbol{A}_m\|$ 为收敛的。

反之，若 $\sum_{m=1}^{\infty}\|\boldsymbol{A}_m\|$ 收敛，则由 $|(\boldsymbol{A}_m)_{ij}|\leqslant\|\boldsymbol{A}_m\|$，$(i,\ j=1,\ 2,\ \cdots,\ n)$，得知在(5)式左边 n^2 个数项级数中每一个级数都是绝对收敛的，故由定义 2 可知级数(4)绝对收敛。

性质 3° 如果 $\sum_{m=1}^{\infty}\boldsymbol{A}_m$ 是收敛或绝对收敛的，那末 $\sum_{m=1}^{\infty}\boldsymbol{P}\boldsymbol{A}_m\boldsymbol{Q}$ 也是收敛或绝对收敛的，并有 $\sum_{m=1}^{\infty}\boldsymbol{P}\boldsymbol{A}_m\boldsymbol{Q}=\boldsymbol{P}\left(\sum_{m=1}^{\infty}\boldsymbol{A}_m\right)\boldsymbol{Q}$.

证明 若 $\sum_{m=1}^{\infty}\boldsymbol{A}_m$ 收敛，令和为 $\boldsymbol{S}$，于是若 $\boldsymbol{S}_N=\sum_{m=1}^{N}\boldsymbol{A}_m$，则 $\lim\limits_{N\to\infty}\boldsymbol{S}_N=\boldsymbol{S}$，根据第 1 节的(3)式，就有：

$$\boldsymbol{P}\boldsymbol{S}_N\boldsymbol{Q}\to\boldsymbol{P}\boldsymbol{S}\boldsymbol{Q},\ \text{当}\ N\to\infty\ \text{时},$$

这就是说 $\sum_{m=1}^{\infty}\boldsymbol{P}\boldsymbol{A}_m\boldsymbol{Q}$ 是收敛的，并且

$$\sum_{m=1}^{\infty} \boldsymbol{P}\boldsymbol{A}_m\boldsymbol{Q}=\boldsymbol{P}\left(\sum_{m=1}^{\infty}\boldsymbol{A}_m\right)\boldsymbol{Q}.$$

再假定 $\sum\limits_{m=1}^{\infty}\boldsymbol{A}_m$ 绝对收敛，则据性质 2° 可知 $\sum\limits_{m=1}^{\infty}\|\boldsymbol{A}_m\|$ 收敛；由前节的性质 1°，又有

$$\|\boldsymbol{P}\boldsymbol{A}_m\boldsymbol{Q}\|\leqslant n^2\|\boldsymbol{P}\|\cdot\|\boldsymbol{A}_m\|\cdot\|\boldsymbol{Q}\|=K\cdot\|\boldsymbol{A}_m\|,$$

而 $K=n^2\|\boldsymbol{P}\|\cdot\|\boldsymbol{Q}\|$ 为与 m 无关的正数，因之 $\sum\limits_{m=1}^{\infty}\|\boldsymbol{P}\boldsymbol{A}_m\boldsymbol{Q}\|$ 收敛，故再据性质 2° 得知 $\sum\limits_{m=1}^{\infty}\boldsymbol{P}\boldsymbol{A}_m\boldsymbol{Q}$ 绝对收敛.

有了上面的这些预备知识，我们就可以讨论矩阵的幂级数了．下面的定理是本节的主要问题，即：

定理 如果 $\boldsymbol{A}$ 之所有特征值都在幂级数

$$\varphi(z)=\sum_{m=0}^{\infty}c_m z^m \tag{6}$$

之收敛圆内，那末矩阵幂级数

$$\sum_{m=0}^{\infty}c_m\boldsymbol{A}^m \tag{7}$$

是绝对收敛的．如果 $\boldsymbol{A}$ 之特征值中有一个在(6)之收敛圆外，那末(7)是发散的.

证明 设

$$\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}=\boldsymbol{J}=\begin{pmatrix}\lambda_1 & & *_1\\ & \lambda_2 & \\ & & \ddots \\ 0 & & & \lambda_n\end{pmatrix}$$

为 $\boldsymbol{A}$ 之约旦标准形，因之 $\lambda_1,\cdots,\lambda_n$ 为 $\boldsymbol{A}$ 的 n 个特征值．因为

$$\boldsymbol{P}^{-1}\boldsymbol{A}^m\boldsymbol{P}=\boldsymbol{J}^m=\begin{pmatrix}\lambda_1^m & & *_m\\ & \lambda_2^m & \\ & & \ddots \\ 0 & & & \lambda_n^m\end{pmatrix},$$

故 $\sum_{m=0}^{N} c_m J^m$ 之主对角线上的元素为 $\sum_{m=0}^{N} c_m \lambda_i^m (i=1, 2, \cdots, n)$.

如有某 $|\lambda_k|>R$, R 为(6)之收敛半径, 则 $\sum_{m=0}^{\infty} c_m \lambda_k^m$ 发散, 因之

$$\sum_{m=0}^{\infty} c_m J^m \tag{8}$$

也发散, 据性质 3° 即知(7)发散.

若每 $|\lambda_i|<R(i=1, 2, \cdots, n)$, 则必可选取两两互异之正实数 $\omega_1, \omega_2, \cdots, \omega_n$, 使

$$|\lambda_i| \leqslant \omega_i < R.$$

再作矩阵

$$T=\begin{pmatrix} \omega_1 & & *_1 \\ & \omega_2 & \\ & & \ddots \\ & & & \omega_n \end{pmatrix},$$

于是 $J \ll T$, 故 $J^m \ll T^m$, $c_m J^m \ll |c_m| \cdot T^m$.

然因有满秩矩阵 Q 使 $Q^{-1}TQ=D=\mathrm{diag}\{\omega_1, \cdots, \omega_n\}$,

故
$$\sum_{m=0}^{N} |c_m| \cdot D^m = \begin{pmatrix} \sum_{m=0}^{N} |c_m| \omega_1^m & & & \\ & \sum_{m=0}^{N} |c_m| \omega_2^m & & \\ & & \ddots & \\ & & & \sum_{m=0}^{N} |c_m| \omega_n^m \end{pmatrix};$$

但因 $0 \leqslant \omega_i < R$, 故 $\sum_{m=0}^{\infty} c_m \omega_i^m$ 绝对收敛, 即 $\sum_{m=0}^{\infty} |c_m| \omega_i^m$ 收敛, 因之 $\lim_{N\to\infty} \sum_{m=0}^{N} |c_m| D^m$ 存在, 于是由性质 3° 知 $\sum_{m=0}^{\infty} |c_m| T^m$ 收敛而有和等于

$$Q\left(\sum_{m=0}^{\infty} |c_m| D^m\right) Q^{-1},$$

再用前节的性质 3°, 又知道

$$\sum_{m=0}^{\infty} c_m \boldsymbol{J}^m$$

收敛. 不仅是收敛, 而且因 $\boldsymbol{J}_1 \ll \boldsymbol{T}$ ($\boldsymbol{J}_1$ 表示把 $\boldsymbol{J}$ 之每元换为其绝对值后所成的非负矩阵), 故 $|c_m|\boldsymbol{J}_1^m \ll |c_m|\boldsymbol{T}^m$, 由是 $\sum\limits_{m=0}^{\infty}|c_m|\boldsymbol{J}_1^m$ 收敛, 此即表示

$$\sum_{m=0}^{\infty} c_m \boldsymbol{J}^m$$

绝对收敛. 故据性质 3°, 知

$$\sum_{m=0}^{\infty} c_m \boldsymbol{A}^m = \sum_{m=0}^{\infty} c_m \boldsymbol{P}\boldsymbol{J}^m \boldsymbol{P}^{-1}$$

为绝对收敛的. 证完.

由这定理即得下述推论:

推论 1 *若幂级数(6)在整个复平面内是收敛的, 则不论 $\boldsymbol{A}$ 为任何矩阵, (7)总是绝对收敛的.*

推论 1 还可直接地简证于下: 因对正整数 m, 可用归纳法, 易证 $\|\boldsymbol{A}^m\| \leqslant (n\cdot\|\boldsymbol{A}\|)^m$. 于是,

$$\sum_{m=0}^{N}\|c_m \boldsymbol{A}^m\| \leqslant \sum_{m=0}^{N}|c_m|\cdot(n\|\boldsymbol{A}\|)^m < K,$$

但 K 与 N 无关. 因之, 据性质 2°, 知(7)绝对收敛.

推论 2 *如果 $\boldsymbol{A}$ 之特征值的模都小于 1, 则级数*

$$\sum_{m=0}^{\infty} \boldsymbol{A}^m = \boldsymbol{E} + \boldsymbol{A} + \boldsymbol{A}^2 + \boldsymbol{A}^3 + \cdots$$

是绝对收敛的, 且和等于 $(\boldsymbol{E}-\boldsymbol{A})^{-1}$.

证明 因为 $\sum\limits_{m=0}^{\infty} z^m$ 的收敛半径为 1, 故由定理即知 $\sum\limits_{m=0}^{\infty} \boldsymbol{A}^m$ 是绝对收敛的. 又由于 $\boldsymbol{A}$ 之每特征值的模都小于 1, 故 $(\boldsymbol{E}-\boldsymbol{A})^{-1}$ 存在; 于是, 若令

$$S_N = E + A + A^2 + \cdots + A^{N-1},$$

则 $S_N \cdot (E - A) = E - A^N$，故

$$S_N = (E - A^N) \cdot (E - A)^{-1} = (E - A)^{-1} - A^N \cdot (E - A)^{-1},$$

因而

$$\lim_{N\to\infty} S_N = (E - A)^{-1} - (\lim_{N\to\infty} A^N) \cdot (E - A)^{-1}$$
$$= (E - A)^{-1} - O \cdot (E - A)^{-1} = (E - A)^{-1}.$$

证完。

注意：推论 2 的逆也为真，即：

推论 3 若 $\sum_{m=0}^{\infty} A^m = E + A + A^2 + A^3 + \cdots$ 收敛，则 A 之特征值的模必都小于 1。

证明 设 $\lambda_1, \lambda_2, \cdots, \lambda_n$ 为 A 之 n 个特征值，而令 $\varphi_m(x) = 1 + x + x^2 + \cdots + x^m$，则 $S_m = E + A + A^2 + \cdots + A^m$ 之 n 个特征值为 $\varphi_m(\lambda_1)$，$\varphi_m(\lambda_2)$，$\cdots$，$\varphi_m(\lambda_n)$。作 A 之约旦标准形

$$P^{-1}AP = \begin{pmatrix} \lambda_1 & & *_1 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$$

后，就应有

$$P^{-1}S_mP = \begin{pmatrix} \varphi_m(\lambda_1) & & *_m \\ & \varphi_m(\lambda_2) & \\ & & \ddots \\ 0 & & & \varphi_m(\lambda_n) \end{pmatrix}.$$

已知 $\sum_{m=0}^{\infty} A^m$ 收敛，意味着 $\lim_{m\to\infty} S_m$ 存在，设为 $\lim_{m\to\infty} S_m = S$，于是，由性质 3° 即得 $\lim_{m\to\infty} P^{-1}S_mP = P^{-1}SP$，因而 $\lim_{m\to\infty} \varphi_m(\lambda_i)$ 均存在 $(i = 1, 2, \cdots, n)$，即序列 $\varphi_1(\lambda_i)$，$\varphi_2(\lambda_i)$，$\cdots$，$\varphi_m(\lambda_i)$，$\cdots$ 有极限，即序列 $1 + \lambda_i$，$1 + \lambda_i + \lambda_i^2$，$\cdots$，$1 + \lambda_i + \lambda_i^2 + \cdots + \lambda_i^m$，$\cdots$ 有极限，或级数 $\sum_{m=0}^{\infty} \lambda_i^m$ 收敛，故有 $|\lambda_i| < 1$，证完。

第 3 节　矩阵的指数函数与三角函数

在函数论中，我们已经知道：

$$e^z = 1 + \frac{z}{1!} + \frac{z^2}{2!} + \frac{z^3}{3!} + \cdots,$$

$$\cos z = 1 - \frac{1}{2!}z^2 + \frac{1}{4!}z^4 - \frac{1}{6!}z^6 + \cdots,$$

$$\sin z = z - \frac{1}{3!}z^3 + \frac{1}{5!}z^5 - \cdots,$$

都在整个平面上是收敛的．于是，据前节定理的推论 1，就知道：不论 $\boldsymbol{A}$ 是任何的矩阵，级数

$$\boldsymbol{E} + \boldsymbol{A} + \frac{1}{2!}\boldsymbol{A}^2 + \frac{1}{3!}\boldsymbol{A}^3 + \cdots,$$

$$\boldsymbol{E} - \frac{1}{2!}\boldsymbol{A}^2 + \frac{1}{4!}\boldsymbol{A}^4 - \cdots,$$

$$\boldsymbol{A} - \frac{1}{3!}\boldsymbol{A}^3 + \frac{1}{5!}\boldsymbol{A}^5 - \cdots$$

都是绝对收敛的，故都有和．由于分析学知识的启发，很自然地用 $e^{\boldsymbol{A}}$，$\cos \boldsymbol{A}$，$\sin \boldsymbol{A}$ 分别表示它们的和，即

$$e^{\boldsymbol{A}} = \boldsymbol{E} + \boldsymbol{A} + \frac{1}{2!}\boldsymbol{A}^2 + \frac{1}{3!}\boldsymbol{A}^3 + \cdots,$$

$$\cos \boldsymbol{A} = \boldsymbol{E} - \frac{1}{2!}\boldsymbol{A}^2 + \frac{1}{4!}\boldsymbol{A}^4 - \cdots,$$

$$\sin \boldsymbol{A} = \boldsymbol{A} - \frac{1}{3!}\boldsymbol{A}^3 + \frac{1}{5!}\boldsymbol{A}^5 - \cdots.$$

这样，利用第 2 节的性质 1°，即可推得

$$\left.\begin{aligned} e^{iA}&=\cos A+i\sin A,\\ \cos A&=\frac{1}{2}(e^{iA}+e^{-iA}),\\ \sin A&=\frac{1}{2i}(e^{iA}-e^{-iA}),\\ \cos(-A)=\cos A,&\ \sin(-A)=-\sin A. \end{aligned}\right\}\qquad(9)$$

至于给定了矩阵 A，怎样实际地计算 e^A、$\cos A$、$\sin A$，其方法虽各有不同，但若能把 A 之特征值求出，往往都易于计算．今举二例说明于下．

［例 1］ 已知四级矩阵 A 之特征值为 π、$-\pi$、0、0，求 $\sin A$、$\cos A$ 及 e^A．

解 因 A 之特征方程为 $|\lambda E-A|=(\lambda-\pi)(\lambda+\pi)\lambda^2=\lambda^4-\pi^2\lambda^2=0$，故 $A^4=\pi^2A^2$，因之，

$$\begin{aligned}\sin A&=A-\frac{1}{3!}A^3+\frac{1}{5!}A^5-\cdots=A-\frac{1}{3!}A^3\\ &\quad+\frac{1}{5!}\pi^2A^3-\frac{1}{7!}\pi^4A^3+\frac{1}{9!}\pi^6A^3-\cdots\\ &=A+\left(-\frac{1}{3!}+\frac{1}{5!}\pi^2-\frac{1}{7!}\pi^4+\frac{1}{9!}\pi^6-\cdots\right)A^3\\ &=A+\frac{\sin\pi-\pi}{\pi^3}A^3=A-\pi^{-2}A^3;\end{aligned}$$

$$\begin{aligned}\cos A&=E-\frac{1}{2!}A^2+\frac{1}{4!}A^4-\frac{1}{6!}A^6+\frac{1}{8!}A^8-\cdots\\ &=E-\frac{1}{2!}A^2+\frac{1}{4!}\pi^2A^2-\frac{1}{6!}\pi^4A^2+\frac{1}{8!}\pi^6A^2-\cdots\\ &=E+\left(-\frac{1}{2!}+\frac{1}{4!}\pi^2-\frac{1}{6!}\pi^4+\frac{1}{8!}\pi^6-\cdots\right)A^2\\ &=E+\frac{\cos\pi-1}{\pi^2}A^2=E-2\pi^{-2}A^2;\end{aligned}$$

$$
\begin{aligned}
e^A &= \boldsymbol{E}+\boldsymbol{A}+\frac{1}{2!}\boldsymbol{A}^2+\frac{1}{3!}\boldsymbol{A}^3+\frac{1}{4!}\boldsymbol{A}^4+\cdots \\
&= \boldsymbol{E}+\boldsymbol{A}+\frac{1}{2!}\boldsymbol{A}^2+\frac{1}{3!}\boldsymbol{A}^3+\frac{1}{4!}\pi^2\boldsymbol{A}^2+\frac{1}{5!}\pi^2\boldsymbol{A}^3 \\
&\quad +\frac{1}{6!}\pi^4\boldsymbol{A}^2+\frac{1}{7!}\pi^4\boldsymbol{A}^3+\cdots \\
&= \boldsymbol{E}+\boldsymbol{A}+\left(\frac{1}{2!}+\frac{1}{4!}\pi^2+\frac{1}{6!}\pi^4+\cdots\right)\boldsymbol{A}^2 \\
&\quad +\left(\frac{1}{3!}+\frac{1}{5!}\pi^2+\frac{1}{7!}\pi^4+\cdots\right)\boldsymbol{A}^3 \\
&= \boldsymbol{E}+\boldsymbol{A}+\frac{p}{\pi^2}\boldsymbol{A}^2+\frac{e^\pi-\pi-1-p}{\pi^3}\boldsymbol{A}^3,
\end{aligned}
$$

这里的 $p=\frac{1}{2!}\pi^2+\frac{1}{4!}\pi^4+\frac{1}{6!}\pi^6+\cdots=\frac{e^\pi+e^{-\pi}}{2}-1$.

[例 2] 已知 $\boldsymbol{A}\sim\operatorname{diag}\{\lambda_1, \lambda_2, \cdots, \lambda_n\}$, 求 e^A, $\cos\boldsymbol{A}$ 与 $\sin\boldsymbol{A}$.

解 因有满秩矩阵 $\boldsymbol{S}$ 使 $\boldsymbol{S}^{-1}\boldsymbol{A}\boldsymbol{S}=\operatorname{diag}\{\lambda_1, \lambda_2, \cdots, \lambda_n\}$, 故

$$\boldsymbol{S}^{-1}\boldsymbol{A}^k\boldsymbol{S}=\operatorname{diag}\{\lambda_1^k, \lambda_2^k, \cdots, \lambda_n^k\},$$

这样, 有 $\boldsymbol{S}^{-1}\cdot f(\boldsymbol{A})\cdot\boldsymbol{S}=\operatorname{diag}\{f(\lambda_1), f(\lambda_2), \cdots, f(\lambda_n)\}$, $f(z)$ 是 z 之多项式. 故若 $\varphi(z)=\sum_{m=0}^{\infty}c_mz^m$ 在全平面内收敛, 则 $\varphi(\boldsymbol{A})=\sum_{m=0}^{\infty}c_m\boldsymbol{A}^m$ 为绝对收敛, 因之, 令 $\varphi_N(\boldsymbol{A})=\sum_{m=0}^{N}c_m\boldsymbol{A}^m$, 则有 $\varphi_N(\boldsymbol{A})\to\varphi(\boldsymbol{A})$ 当 $N\to\infty$ 时. 再将上面的 $f(z)$ 用 $\varphi_N(z)=\sum_{m=0}^{N}c_mz^m$ 代换之, 即得

$$\boldsymbol{S}^{-1}\cdot\varphi_N(\boldsymbol{A})\cdot\boldsymbol{S}=\operatorname{diag}\{\varphi_N(\lambda_1), \varphi_N(\lambda_2), \cdots, \varphi_N(\lambda_n)\};$$

于是, 从 $\lim_{N\to\infty}\varphi_N(\lambda_i)\to\varphi(\lambda_i)=\sum_{m=0}^{\infty}c_m\lambda_i^m$ 及 $\lim_{N\to\infty}\varphi_N(\boldsymbol{A})\to\varphi(\boldsymbol{A})$,

则得

$$S^{-1}\cdot\varphi(A)\cdot S=\mathrm{diag}\{\varphi(\lambda_1),\ \varphi(\lambda_2),\ \cdots,\ \varphi(\lambda_n)\}.$$

故若将 $\varphi(z)$ 分别以 e^z, $\sin z$, $\cos z$ 代入之，就有

$$e^A=S\cdot\mathrm{diag}\{e^{\lambda_1},\ e^{\lambda_2},\ \cdots,\ e^{\lambda_n}\}\cdot S^{-1},$$

$$\sin A=S\cdot\mathrm{diag}\{\sin\lambda_1,\ \sin\lambda_2,\ \cdots,\ \sin\lambda_n\}\cdot S^{-1},$$

$$\cos A=S\cdot\mathrm{diag}\{\cos\lambda_1,\ \cos\lambda_2,\ \cdots,\ \cos\lambda_n\}\cdot S^{-1}.$$

从例 2 的启发，易知 e^A、$\sin A$、$\cos A$ 之特征值与 A 之特征值间的关系．即有：

定理 1 设 A 之特征值为 λ_1、λ_2、$\cdots$、λ_n，则：

i) e^A 之特征值为 e^{λ_1}, e^{λ_2}, $\cdots$, e^{λ_n};

ii) $\sin A$ 之特征值为 $\sin\lambda_1$, $\sin\lambda_2$, $\cdots$, $\sin\lambda_n$;

iii) $\cos A$ 之特征值为 $\cos\lambda_1$, $\cos\lambda_2$, $\cdots$, $\cos\lambda_n$.

证明 令

$$P^{-1}AP=J=\begin{pmatrix}\lambda_1 & & *_1\\ & \lambda_2 & \\ & & \ddots \\ 0 & & \lambda_n\end{pmatrix}$$

为 A 之约旦标准形．由前节的性质 3°，有

$$e^J=P^{-1}\cdot e^A\cdot P,$$

即 $e^J\sim e^A$．同理，也有 $\sin J\sim\sin A$, $\cos J\sim\cos A$．故 e^A 之特征值与 e^J 同，于是只须求出 e^J 的特征值就行了．对 $\sin A$ 与 $\cos A$，也类似．因为

$$e^J=E+J+\frac{1}{2!}J^2+\frac{1}{3!}J^3+\cdots=\sum_{m=0}^{\infty}\frac{1}{m!}J^m,$$

故令 $S_N=\sum_{m=0}^{N}\frac{1}{m!}J^m$ 时，则因每 $J^m=\begin{pmatrix}\lambda_1^m & & *_m\\ & \lambda_2^m & \\ & & \ddots \\ 0 & & \lambda_n^m\end{pmatrix}$ 均为

(上)三角矩阵，故 $\boldsymbol{S}_N$ 亦必是上三角矩阵：

$$\boldsymbol{S}_N=\begin{pmatrix}\sum_{m=0}^{N}\frac{1}{m!}\lambda_1^m & & * \\ & \sum_{m=0}^{N}\frac{1}{m!}\lambda_2^m & \\ & & \ddots \\ 0 & & & \sum_{m=0}^{N}\frac{1}{m!}\lambda_n^m\end{pmatrix},$$

因之

$$e^{J}=\lim_{N\to\infty}\boldsymbol{S}_N=\begin{pmatrix}\lim\limits_{N\to\infty}\sum_{m=0}^{N}\frac{1}{m!}\lambda_1^m & & *' \\ & \lim\limits_{N\to\infty}\sum_{m=0}^{N}\frac{1}{m!}\lambda_2^m & \\ & & \ddots \\ 0 & & & \lim\limits_{N\to\infty}\sum_{m=0}^{N}\frac{1}{m!}\lambda_n^m\end{pmatrix}$$

$$=\begin{pmatrix}e^{\lambda_1} & & *' \\ & e^{\lambda_2} & \\ & & \ddots \\ 0 & & & e^{\lambda_n}\end{pmatrix},$$

证明了 $e^{J}(\sim e^{A})$ 之全部特征值为 $e^{\lambda_1}, e^{\lambda_2}, \cdots, e^{\lambda_n}$. 同样的方法，可证 ii) 与 iii). 定理 1 证完.

我们知道，在分析学中，指数函数 e^z 有一个重要性质，即 $e^{z_1+z_2}=e^{z_1}e^{z_2}=e^{z_2}e^{z_1}$. 在矩阵分析中，这样的关系不再成立了. 例如，设

$$\boldsymbol{A}=\begin{pmatrix}1 & 1\\ 0 & 0\end{pmatrix},\quad \boldsymbol{B}=\begin{pmatrix}1 & -1\\ 0 & 0\end{pmatrix},$$

则 $\boldsymbol{A}^2=\boldsymbol{A}$, $\boldsymbol{B}^2=\boldsymbol{B}$, 故

$$\boldsymbol{A}=\boldsymbol{A}^2=\boldsymbol{A}^3=\boldsymbol{A}^4=\cdots,\qquad \boldsymbol{B}=\boldsymbol{B}^2=\boldsymbol{B}^3=\boldsymbol{B}^4=\cdots,$$

于是
$$e^{A}=\boldsymbol{E}+(e-1)\boldsymbol{A}=\begin{pmatrix} e & e-1 \\ 0 & 1 \end{pmatrix},$$

$$e^{B}=\boldsymbol{E}+(e-1)\boldsymbol{B}=\begin{pmatrix} e & 1-e \\ 0 & 1 \end{pmatrix},$$

因之
$$e^{\boldsymbol{A}}\cdot e^{\boldsymbol{B}}=\begin{pmatrix} e^2 & -(e-1)^2 \\ 0 & 1 \end{pmatrix},$$

$$e^{\boldsymbol{B}}\cdot e^{\boldsymbol{A}}=\begin{pmatrix} e^2 & (e-1)^2 \\ 0 & 1 \end{pmatrix}.$$

又由
$$\boldsymbol{A}+\boldsymbol{B}=\begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix}$$

得
$$(\boldsymbol{A}+\boldsymbol{B})^2=2(\boldsymbol{A}+\boldsymbol{B}),$$

故归纳地可证 $(\boldsymbol{A}+\boldsymbol{B})^k=2^{k-1}\cdot(\boldsymbol{A}+\boldsymbol{B})$, 由是易知

$$e^{\boldsymbol{A}+\boldsymbol{B}}=\boldsymbol{E}+\frac{1}{2}(e^2-1)(\boldsymbol{A}+\boldsymbol{B})=\begin{pmatrix} e^2 & 0 \\ 0 & 1 \end{pmatrix},$$

这个例就已说明了确有可能使 $e^{\boldsymbol{A}+\boldsymbol{B}}$、$e^{\boldsymbol{A}}\cdot e^{\boldsymbol{B}}$、$e^{\boldsymbol{B}}\cdot e^{\boldsymbol{A}}$ 中两两互不相等, 即关系式

$$e^{\boldsymbol{A}+\boldsymbol{B}}=e^{\boldsymbol{A}}\cdot e^{\boldsymbol{B}}=e^{\boldsymbol{B}}\cdot e^{\boldsymbol{A}}$$

不再普遍成立. 可是, 当 $\boldsymbol{AB}=\boldsymbol{BA}$ 时, 上关系式还是成立的. 为此, 需要先引进矩阵的微分及积分的概念.

定义 1 若矩阵 $\boldsymbol{A}(t)=(a_{ij}(t))$之每元 $a_{ij}(t)$为变量 t 的可微函数, 就叫矩阵 $A(t)$是可微的, 其导数(微商)用公式

$$\dot{\boldsymbol{A}}(t)=\frac{d}{dt}\boldsymbol{A}(t)=\left(\frac{d}{dt}a_{ij}(t)\right)$$

来定义. 同样, 用公式

$$\int_{\alpha}^{\beta}\boldsymbol{A}(t)dt=\left(\int_{\alpha}^{\beta}a_{ij}(t)dt\right)$$

来定义矩阵 $\boldsymbol{A}(t)$ 的定积分. 至于矩阵 $\boldsymbol{A}(t)$ 叫做可积, 连续, 或有界, 指的是每元 $a_{ij}(t)$ 都具有这类似的性质.

定理 2 设 $\boldsymbol{A}(t)$ 与 $\boldsymbol{B}(t)$ 是两个可微矩阵, 且积 $\boldsymbol{A}(t)\boldsymbol{B}(t)$ 有意义, 那末

$$\frac{d}{dt}\{\boldsymbol{A}(t)\boldsymbol{B}(t)\}=\dot{\boldsymbol{A}}(t)\cdot\boldsymbol{B}(t)+\boldsymbol{A}(t)\cdot\dot{\boldsymbol{B}}(t).$$

证明 令 $\boldsymbol{A}(t)\cdot\boldsymbol{B}(t)=\boldsymbol{P}(t)=(p_{ij}(t))$, 则

$$p_{ij}(t)=\sum_r a_{ir}(t)\cdot b_{rj}(t),$$

故

$$\frac{d}{dt}p_{ij}(t)=\sum_r\left\{\frac{d}{dt}a_{ir}(t)\right\}\cdot b_{rj}(t)+\sum_r a_{ir}(t)\cdot\left\{\frac{d}{dt}b_{rj}(t)\right\},$$

这就是说:

$$\frac{d}{dt}\boldsymbol{P}(t)=\dot{\boldsymbol{A}}(t)\cdot\boldsymbol{B}(t)+\boldsymbol{A}(t)\cdot\dot{\boldsymbol{B}}(t).$$

关于矩阵的指数函数, 有:

定理 3 不论 $\boldsymbol{A}$ 是任何矩阵, 常有

$$\frac{d}{dt}e^{t\boldsymbol{A}}=\boldsymbol{A}\cdot e^{t\boldsymbol{A}}=e^{t\boldsymbol{A}}\cdot\boldsymbol{A}.$$

证明 首先要注意的是

$$(e^{t\boldsymbol{A}})_{ij}=\sum_{m=0}^{\infty}\frac{1}{m!}t^m\cdot(\boldsymbol{A}^m)_{ij},$$

但上式右边是 t 的幂级数, 不管 t 为何值, 它总是收敛的, 因之可逐项微分, 有

$$\frac{d}{dt}\{(e^{t\boldsymbol{A}})_{ij}\}=\sum_{m=1}^{\infty}\frac{1}{(m-1)!}t^{m-1}\cdot(\boldsymbol{A}^m)_{ij},$$

于是, 由第 2 节之性质 3°, 就有

$$\frac{d}{dt}e^{tA}=\sum_{m=1}^{\infty}\frac{1}{(m-1)!}t^{m-1}\boldsymbol{A}^m$$

$$=\begin{cases}\boldsymbol{A}\cdot\sum_{m=1}^{\infty}\frac{1}{(m-1)!}t^{m-1}\boldsymbol{A}^{m-1}=\boldsymbol{A}\cdot e^{t\boldsymbol{A}},\\ \left\{\sum_{m=1}^{\infty}\frac{1}{(m-1)!}t^{m-1}\boldsymbol{A}^{m-1}\right\}\cdot\boldsymbol{A}=e^{t\boldsymbol{A}}\cdot\boldsymbol{A}.\end{cases}$$

现在可证明我们所需要的结果，即：

定理 4 如果 $\boldsymbol{AB}=\boldsymbol{BA}$，则

$$e^{\boldsymbol{A}}\cdot e^{\boldsymbol{B}}=e^{\boldsymbol{B}}\cdot e^{\boldsymbol{A}}=e^{\boldsymbol{A}+\boldsymbol{B}}.$$

证明 由前节性质 3°，有

$$e^{t\boldsymbol{A}}\cdot\boldsymbol{B}=\left(\sum_{m=0}^{\infty}\frac{t^m}{m!}\boldsymbol{A}^m\right)\cdot\boldsymbol{B}$$

$$=\sum_{m=0}^{\infty}\frac{t^m}{m!}\boldsymbol{A}^m\boldsymbol{B}=\sum_{m=0}^{\infty}\frac{t^m}{m!}\boldsymbol{B}\boldsymbol{A}^m$$

$$=\boldsymbol{B}\cdot\sum_{m=0}^{\infty}\frac{t^m}{m!}\boldsymbol{A}^m=\boldsymbol{B}\cdot e^{t\boldsymbol{A}};$$

即当 $\boldsymbol{AB}=\boldsymbol{BA}$ 时，有

$$e^{t\boldsymbol{A}}\cdot\boldsymbol{B}=\boldsymbol{B}\cdot e^{t\boldsymbol{A}};\tag{10}$$

同理，又有

$$e^{t\boldsymbol{B}}\cdot\boldsymbol{A}=\boldsymbol{A}\cdot e^{t\boldsymbol{B}}.\tag{11}$$

今令 $\boldsymbol{C}(t)=e^{t(\boldsymbol{A}+\boldsymbol{B})}\cdot e^{-t\boldsymbol{A}}\cdot e^{-t\boldsymbol{B}}$，则由定理 2 及 3 可得

$$\dot{\boldsymbol{C}}(t)=(\boldsymbol{A}+\boldsymbol{B})\cdot e^{t(\boldsymbol{A}+\boldsymbol{B})}\cdot e^{-t\boldsymbol{A}}\cdot e^{-t\boldsymbol{B}}$$

$$+e^{t(\boldsymbol{A}+\boldsymbol{B})}\cdot(-\boldsymbol{A})e^{-t\boldsymbol{A}}\cdot e^{-t\boldsymbol{B}}+e^{t(\boldsymbol{A}+\boldsymbol{B})}\cdot e^{-t\boldsymbol{A}}\cdot(-\boldsymbol{B})e^{-t\boldsymbol{B}};$$

因 $\boldsymbol{AB}=\boldsymbol{BA}$，故 $\boldsymbol{A}+\boldsymbol{B}$ 与 $\boldsymbol{A}$ 及 $\boldsymbol{B}$ 均可换（关于乘法），于是由(10)及(11)则知上式的右边等于

$$(\boldsymbol{A}+\boldsymbol{B}-\boldsymbol{A}-\boldsymbol{B})e^{t(\boldsymbol{A}+\boldsymbol{B})}\cdot e^{-t\boldsymbol{A}}\cdot e^{-t\boldsymbol{B}}=\boldsymbol{0},$$

因之
$$\dot{\boldsymbol{C}}(t)=\frac{d}{dt}\boldsymbol{C}(t)=\boldsymbol{0},$$

这说明了 $\boldsymbol{C}(t)$ 实际上与 t 无关，故有

$$\boldsymbol{C}(1)=\boldsymbol{C}(0).$$

但 $\boldsymbol{C}(0)=\boldsymbol{E}$, 故 $\boldsymbol{C}(1)=\boldsymbol{E}$, 即

$$e^{\boldsymbol{A}+\boldsymbol{B}}\cdot e^{-\boldsymbol{A}}\cdot e^{-\boldsymbol{B}}=\boldsymbol{E}. \tag{12}$$

同理(据 $\boldsymbol{A}$, $\boldsymbol{B}$ 之位置的对称原故), 又有

$$e^{\boldsymbol{A}+\boldsymbol{B}}\cdot e^{-\boldsymbol{B}}\cdot e^{-\boldsymbol{A}}=\boldsymbol{E}. \tag{13}$$

这就表示: 当 $\boldsymbol{AE}=\boldsymbol{BA}$ 时, 有(12)与(13)二个关系式. 因而, 特取 $\boldsymbol{B}=-\boldsymbol{A}$, 由(12)及(13)即得

$$e^{0}\cdot e^{-\boldsymbol{A}}\cdot e^{\boldsymbol{A}}=\boldsymbol{E} \quad 及 \quad e^{0}\cdot e^{\boldsymbol{A}}\cdot e^{-\boldsymbol{A}}=\boldsymbol{E}.$$

但 $e^{0}=\boldsymbol{E}$ 是显然的, 故对于任何矩阵 $\boldsymbol{A}$ 恒有

$$e^{-\boldsymbol{A}}\cdot e^{\boldsymbol{A}}=\boldsymbol{E}=e^{\boldsymbol{A}}\cdot e^{-\boldsymbol{A}},$$

即 $e^{-\boldsymbol{A}}=(e^{\boldsymbol{A}})^{-1}$. 于是(12)与(13)就分别变成了

$$e^{\boldsymbol{A}+\boldsymbol{B}}\cdot(e^{\boldsymbol{A}})^{-1}\cdot(e^{\boldsymbol{B}})^{-1}=\boldsymbol{E}$$

与

$$e^{\boldsymbol{A}+\boldsymbol{B}}\cdot(e^{\boldsymbol{B}})^{-1}\cdot(e^{\boldsymbol{A}})^{-1}=\boldsymbol{E},$$

亦即

$$e^{\boldsymbol{A}+\boldsymbol{B}}=e^{\boldsymbol{B}}\cdot e^{\boldsymbol{A}}=e^{\boldsymbol{A}}\cdot e^{\boldsymbol{B}},$$

证完.

在证定理 4 的过程中, 已经证得了下面的:

定理 5 不论 $\boldsymbol{A}$ 是任何矩阵, $e^{\boldsymbol{A}}$ 总是满秩的, 且 $(e^{\boldsymbol{A}})^{-1}=e^{-\boldsymbol{A}}$.

利用定理 4, 还可以推导一些有趣的结果. 例如, 当 $\boldsymbol{AB}=\boldsymbol{BA}$ 时, 从公式(9)以及定理 4, 可知

$$\begin{aligned}\cos(\boldsymbol{A}+\boldsymbol{B})&=\frac{1}{2}\{e^{i(\boldsymbol{A}+\boldsymbol{B})}+e^{-i(\boldsymbol{A}+\boldsymbol{B})}\}=\frac{1}{2}\{e^{i\boldsymbol{A}}\cdot e^{i\boldsymbol{B}}+e^{-i\boldsymbol{A}}\cdot e^{-i\boldsymbol{B}}\}\\&=\frac{1}{2}\left\{\frac{(e^{i\boldsymbol{A}}+e^{-i\boldsymbol{A}})(e^{i\boldsymbol{B}}+e^{-i\boldsymbol{B}})}{2}\right.\\&\qquad\left.+\frac{(e^{i\boldsymbol{A}}-e^{-i\boldsymbol{A}})(e^{i\boldsymbol{B}}-e^{-i\boldsymbol{B}})}{2}\right\}\\&=\frac{e^{i\boldsymbol{A}}+e^{-i\boldsymbol{A}}}{2}\cdot\frac{e^{i\boldsymbol{B}}+e^{-i\boldsymbol{B}}}{2}-\frac{e^{i\boldsymbol{A}}-e^{-i\boldsymbol{A}}}{2i}\cdot\frac{e^{i\boldsymbol{B}}-e^{-i\boldsymbol{B}}}{2i}\\&=\cos\boldsymbol{A}\cdot\cos\boldsymbol{B}-\sin\boldsymbol{A}\cdot\sin\boldsymbol{B}.\end{aligned}$$

特别地，取 $\boldsymbol{B}=\boldsymbol{A}$ 时，就有 $\cos 2\boldsymbol{A}=\cos^2\boldsymbol{A}-\sin^2\boldsymbol{B}$.

同样，当 $\boldsymbol{AB}=\boldsymbol{BA}$ 时又有 $\sin(\boldsymbol{A}+\boldsymbol{B})=\sin\boldsymbol{A}\cos\boldsymbol{B}+\cos\boldsymbol{A}\sin\boldsymbol{B}$，故有 $\sin 2\boldsymbol{A}=2\sin\boldsymbol{A}\cos\boldsymbol{A}$.

最后，不论 $\boldsymbol{A}$ 为任何的矩阵，定理 5 是说 $e^{\boldsymbol{A}}$ 恒为满秩的. 至于 $\sin\boldsymbol{A}$ 与 $\cos\boldsymbol{A}$ 也常为满秩吗？这只要看一个例子就足以说明答案是否定的. 例如，本节的例 2 中，当某一个 λ（如 λ_1）$=\pi$ 时（虽 $\boldsymbol{A}$ 为满秩，即无一个 λ 为零），则 $\sin\boldsymbol{A}=\boldsymbol{S}\cdot\mathrm{diag}\{0,\ \sin\lambda_2,\ \cdots,\ \sin\lambda_n\}\cdot\boldsymbol{S}^{-1}$ 显然是降秩的. 同理，当 $\lambda_1=\dfrac{\pi}{2}$ 时，$\cos\boldsymbol{A}$ 也是降秩的.

练习三十九

1. 试证：不论 $\boldsymbol{A}$ 为任何矩阵，常有 $\sin^2\boldsymbol{A}+\cos^2\boldsymbol{A}=\boldsymbol{E}$，$e^{\boldsymbol{A}+2\pi i\boldsymbol{E}}=e^{\boldsymbol{A}}$，$\sin(\boldsymbol{A}+2\pi\boldsymbol{E})=\sin\boldsymbol{A}$.
2. 试证：若 $\boldsymbol{A}$ 是实反对称（反厄米特）矩阵，则 $e^{\boldsymbol{A}}$ 为实直交（酉）矩阵.
3. 若 $\boldsymbol{A}$ 是厄米特矩阵，则 $e^{i\boldsymbol{A}}$ 是酉矩阵. 并说明这结论在 $n=1$ 时的意义.
4. 证明关系式

$$\frac{d}{dt}\{\boldsymbol{A}(t)\}^m=m\{\boldsymbol{A}(t)\}^{m-1}\cdot\frac{d}{dt}\boldsymbol{A}(t)$$

一般不成立. 又，在什么情况时，它才成立呢？

5. 设 $\boldsymbol{A}=\begin{pmatrix}0&1&0\\0&0&1\\1&0&0\end{pmatrix}$，求 $e^{\boldsymbol{A}}$，$\sin\boldsymbol{A}$ 及 $\cos\boldsymbol{A}$.

第 4 节　在微分方程组中的应用

本节介绍怎样利用矩阵函数来解一阶线性微分方程组. 先研究有常系数且齐次的特例，再研究具变系数且非齐次的一般情况.

设有一阶线性齐次（具常系数的）微分方程组

$$
\begin{cases}
\dfrac{dx_1}{dt}=a_{11}x_1+a_{12}x_2+\cdots+a_{1n}x_n,\\
\dfrac{dx_2}{dt}=a_{21}x_1+a_{22}x_2+\cdots+a_{2n}x_n,\\
\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\
\dfrac{dx_n}{dt}=a_{n1}x_1+a_{n2}x_2+\cdots+a_{nn}x_n,
\end{cases}
\tag{14}
$$

式中 t 是自变量，$x_i=x_i(t)$ 是 t 的函数$(i=1,\ 2,\ \cdots,\ n)$，a_{ij} 是复数$(i,\ j=1,\ 2,\ \cdots,\ n)$.

将(14)式改写为矩阵方程的形式，如

$$
\dot{\boldsymbol{X}}=\frac{d\boldsymbol{X}}{dt}=\boldsymbol{A}\boldsymbol{X},\tag{15}
$$

这里 $\boldsymbol{A}=\begin{pmatrix}a_{11}&\cdots&a_{1n}\\ \vdots&&\vdots\\ a_{n1}&\cdots&a_{nn}\end{pmatrix}=(a_{ij})$，$\boldsymbol{X}=\begin{pmatrix}x_1\\ x_2\\ \vdots\\ x_n\end{pmatrix}$.

今假定(14)即(15)满足初始条件 $x_i(0)=c_i(i=1,\cdots,\ n)$.

并令 $\boldsymbol{X}_0=\begin{pmatrix}c_1\\ c_2\\ \vdots\\ c_n\end{pmatrix}$. 将每 $x_i(t)$ 展开为马克洛林级数

$$
x_i(t)=x_i(0)+\dot{x}_i(0)\cdot t+\ddot{x}_i(0)\cdot\frac{t^2}{2}+\cdots,
$$

式中 $\dot{x}_i(0)=\left.\dfrac{dx_i}{dt}\right|_{t=0}$，$\ddot{x}_i(0)=\left.\dfrac{d^2x_i}{dt^2}\right|_{t=0}$，$\cdots$. 因之就有

$$
\boldsymbol{X}=\boldsymbol{X}_0+\dot{\boldsymbol{X}}_0\cdot t+\ddot{\boldsymbol{X}}_0\cdot\frac{t^2}{2}+\cdots,\tag{16}
$$

式中 $\dot{\boldsymbol{X}}_0=\left.\dfrac{d\boldsymbol{X}}{dt}\right|_{t=0}$，$\ddot{\boldsymbol{X}}_0=\left.\dfrac{d^2\boldsymbol{X}}{dt^2}\right|_{t=0}$，$\cdots$.

但从(15)式有

$$\frac{d^2\boldsymbol{X}}{dt^2}=\boldsymbol{A}\frac{d\boldsymbol{X}}{dt}=\boldsymbol{A}^2\boldsymbol{X},\quad \frac{d^3\boldsymbol{X}}{dt^3}=\boldsymbol{A}^2\frac{d\boldsymbol{X}}{dt}=\boldsymbol{A}^3\boldsymbol{X},\ \cdots,$$

于是 $\dot{\boldsymbol{X}}_0=\boldsymbol{A}\boldsymbol{X}_0,\ \ddot{\boldsymbol{X}}_0=\boldsymbol{A}^2\boldsymbol{X}_0,\ \dddot{\boldsymbol{X}}_0=\boldsymbol{A}^3\boldsymbol{X}_0,\ \cdots,$

故由(16)即得

$$\begin{aligned}\boldsymbol{X}&=\boldsymbol{X}_0+\dot{\boldsymbol{X}}_0t+\ddot{\boldsymbol{X}}_0\frac{t^2}{2}+\cdots\\&=\boldsymbol{X}_0+t\boldsymbol{A}\boldsymbol{X}_0+\frac{t^2}{2!}\boldsymbol{A}^2\boldsymbol{X}_0+\cdots=e^{t\boldsymbol{A}}\cdot\boldsymbol{X}_0,\end{aligned}\tag{17}$$

这是说：(15)的解一定是(17)之形.

反之，不难证明，(17)确为(15)之解. 事实上，从(17)式而利用前节的定理 2 与 3，就得到

$$\begin{aligned}\dot{\boldsymbol{X}}&=\left(\frac{d}{dt}e^{t\boldsymbol{A}}\right)\cdot\boldsymbol{X}_0+e^{t\boldsymbol{A}}\cdot\frac{d}{dt}\boldsymbol{X}_0\\&=\left(\frac{d}{dt}e^{t\boldsymbol{A}}\right)\cdot\boldsymbol{X}_0=\boldsymbol{A}\cdot e^{t\boldsymbol{A}}\boldsymbol{X}_0=\boldsymbol{A}\boldsymbol{X}.\end{aligned}$$

于是证得了：

定理 1 满足初始条件 $x_1(0)=c_1,\ \cdots,\ x_n(0)=c_n$ 的一阶线性齐次且有常系数的微分方程组

$$\dot{\boldsymbol{X}}=\frac{d\boldsymbol{X}}{dt}=\boldsymbol{A}\boldsymbol{X}$$

有且只有唯一个解：$\boldsymbol{X}=e^{t\boldsymbol{A}}\boldsymbol{X}_0$，这里 $\boldsymbol{A}=(a_{ij})$，

$$\boldsymbol{X}=\begin{pmatrix}x_1(t)\\ \vdots\\ x_n(t)\end{pmatrix},\quad \boldsymbol{X}_0=\begin{pmatrix}c_1\\ \vdots\\ c_n\end{pmatrix}.$$

我们知道满足初始条件 $\boldsymbol{X}(0)=c$ 且只含变量 t 的一个函数 $\boldsymbol{X}(t)$ 的微分方程 $\frac{d}{dt}\boldsymbol{X}(t)=a\cdot\boldsymbol{X}(t)$ 的解是 $\boldsymbol{X}(t)=ce^{at}$. 因之定理 1 恰好是它的推广.

[例 1] 解 $\begin{cases}\dfrac{dx_1}{dt}= & -x_2+x_3,\\ \dfrac{dx_2}{dt}= x_1 & +x_3,\\ \dfrac{dx_3}{dt}=-x_1 & -x_2.\end{cases}$

解 这里，$\boldsymbol{A}=\begin{pmatrix}0&-1&1\\1&0&1\\-1&-1&0\end{pmatrix}$，$|\lambda\boldsymbol{E}-\boldsymbol{A}|=\lambda(\lambda^2+3)$，

即 $\boldsymbol{A}$ 之特征值为 0，$\pm\sqrt{3}i$. 再求出分别对应于特征值 0，$\sqrt{3}i$，$-\sqrt{3}i$ 的三个线性无关的特征向量，它们分别为

$$\begin{pmatrix}-1\\1\\1\end{pmatrix},\ \begin{pmatrix}-\omega\\\omega^2\\1\end{pmatrix},\ \begin{pmatrix}-\omega^2\\\omega\\1\end{pmatrix},$$

其中 $\omega=\frac{1}{2}(-1+\sqrt{3}i)$ 为 1 之一个虚立方根. 于是

$$\boldsymbol{P}=\begin{pmatrix}-1&-\omega&-\omega^2\\1&\omega^2&\omega\\1&1&1\end{pmatrix}$$

为满秩，且有 $t\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}=\mathrm{diag}\{0,\ \sqrt{3}it,\ -\sqrt{3}it\}$. 故由第 3 节的例 2，有

$$e^{tA}=\boldsymbol{P}\cdot\mathrm{diag}\{1,\ e^{\sqrt{3}it},\ e^{-\sqrt{3}it}\}\cdot\boldsymbol{P}^{-1}.$$

只须把右边矩阵具体地写出来，就求得解答为

$$\boldsymbol{X}=\begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix}=e^{tA}\cdot\begin{pmatrix}c_1\\c_2\\c_3\end{pmatrix},$$

这里 c_1，c_2，c_3 是独立参数，而 x_1，x_2，x_3 都是 t 之函数.

[例 2] 解 $\begin{cases} \dfrac{dx_1}{dt} = x_2, \\ \dfrac{dx_2}{dt} = -x_1. \end{cases}$

解 这里，$\boldsymbol{A}=\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$，故 $\boldsymbol{A}^2=-\boldsymbol{E}$，$\boldsymbol{A}^3=-\boldsymbol{A}$，$\boldsymbol{A}^4=\boldsymbol{E}$，$\boldsymbol{A}^5=\boldsymbol{A}$，$\boldsymbol{A}^6=-\boldsymbol{E}$，$\boldsymbol{A}^7=-\boldsymbol{A}$，$\boldsymbol{A}^8=\boldsymbol{E}$，…，而一般有

$$\boldsymbol{A}^{4k}=\boldsymbol{E},\ \boldsymbol{A}^{4k+1}=\boldsymbol{A},\ \boldsymbol{A}^{4k+2}=-\boldsymbol{E},\ \boldsymbol{A}^{4k+3}=-\boldsymbol{A}.$$

故

$$\begin{aligned} e^{tA} &= \left(1-\frac{1}{2!}t^2+\frac{1}{4!}t^4-\frac{1}{6!}t^6+\cdots\right)\boldsymbol{E} \\ &\quad +\left(t-\frac{1}{3!}t^3+\frac{1}{5!}t^5-\frac{1}{7!}t^7+\cdots\right)\boldsymbol{A} \\ &= \cos t\cdot\boldsymbol{E}+\sin t\cdot\boldsymbol{A}=\begin{pmatrix} \cos t & \sin t \\ -\sin t & \cos t \end{pmatrix}, \end{aligned}$$

于是，$$\boldsymbol{X}=\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}=e^{tA}\cdot\boldsymbol{X}_0=\begin{pmatrix} \cos t & \sin t \\ -\sin t & \cos t \end{pmatrix}\begin{pmatrix} c_1 \\ c_2 \end{pmatrix},$$

即 $$\begin{cases} x_1=c_1\cos t+c_2\sin t, \\ x_2=-c_1\sin t+c_2\cos t. \end{cases}$$

再研究具变系数的一阶线性非齐次微分方程组：

$$\begin{cases} \dfrac{dx_1}{dt}=a_{11}(t)\cdot x_1+\cdots+a_{1n}(t)\cdot x_n+b_1(t), \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ \dfrac{dx_n}{dt}=a_{n1}(t)\cdot x_1+\cdots+a_{nn}(t)\cdot x_n+b_n(t), \end{cases} \tag{18}$$

式中 $a_{ij}(t)$、$b_i(t)$ 都是定义在区间$[\alpha, \beta]$上的 t 之连续函数，并假定(18)满足初始条件 $x_1(0)=c_1$，…，$x_n(0)=c_n$，——当然已默许了区间$[\alpha, \beta]$含 0。

我们要证明(18)有唯一组解：$x_1(t)$, $\cdots$, $x_n(t)$，当然要求 $x_1(t)$, $\cdots$, $x_n(t)$ 都是定义在区间$[\alpha, \beta]$上连续且可微的函数。

若令 $\boldsymbol{A}(t)=(a_{ij}(t))$为 n 级矩阵，$\boldsymbol{B}(t)=\begin{pmatrix} b_1(t) \\ \vdots \\ b_n(t) \end{pmatrix}$与 $\boldsymbol{X}=\boldsymbol{X}(t)=\begin{pmatrix} x_1(t) \\ \vdots \\ x_n(t) \end{pmatrix}$均为 n 元列向量，则(18)式可以简写成矩阵的形式

$$\dot{\boldsymbol{X}}=\frac{d\boldsymbol{X}}{dt}=\boldsymbol{A}(t)\cdot\boldsymbol{X}+\boldsymbol{B}(t). \tag{19}$$

于是问题就变成了要证明(19)式有且只有唯一的解向量 $\boldsymbol{X}(t)$。

未证明之前，先回忆一下，在微分方程教程中，一阶线性微分方程

$$\frac{dx}{dt}=a(t)\cdot x+b(t) \tag{20}$$

有一个积分因子 $f(t)=e^{-\int a(t)dt}$，从而可导出(20)的解为

$$x=e^{\int a(t)dt}\cdot\int e^{-\int a(t)dt}\cdot b(t)dt=f(t)^{-1}\cdot\int f(t)b(t)dt. \tag{21}$$

于是，由(21)，很容易猜测(20)之推广式(19)的解为

$$\boldsymbol{X}=e^{\int \boldsymbol{A}(t)dt}\cdot\int e^{-\int \boldsymbol{A}(t)dt}\cdot\boldsymbol{B}(t)dt.$$

但我们不采取这样的思路，这是因为：$e^{\boldsymbol{A}}$ 定义中，$\boldsymbol{A}$ 是数量矩阵，因而 $\boldsymbol{A}$ 之特征值全在复数域内；当 $\boldsymbol{A}$ 已给时，其特征值当然全在整个复平面内，从而当然在幂级数 $e^z=1+z+\frac{1}{2!}z^2+\frac{1}{3!}z^3+\cdots$ 的收敛圆内（其收敛半径为 ∞），于是 $\boldsymbol{E}+\boldsymbol{A}+$

$\frac{1}{2!}A^2+\frac{1}{3!}A^3+\cdots$ 收敛而记成 e^A. 根据这个思想, e^{tA}(t 为变量)也是有意义的, 因为 t 虽是变量, 但 tA 之特征值为 $t\lambda_1$, $t\lambda_2$, $\cdots$, $t\lambda_n$(其中 λ_1, $\cdots$, λ_n 为 A之特征值), 故可把 t 看作参变数, 而知 tA 之特征值 $t\lambda_1$, $\cdots$, $t\lambda_n$ 也全在复平面内, 随而可定义 e^{tA}, 即 $E+tA+\frac{1}{2!}t^2A^2+\frac{1}{3!}t^3A^3+\cdots$收敛, 定义其和为 e^{tA}. 可是, 当 $B(t)=(b_{ij}(t))$ 为 n 级连续矩阵时, 其特征值一般为变量 t 之函数, 已不在复数域内, 而应在由一切连续函数所成之函数域的闭扩张域内, 但 $1+z+\frac{1}{2!}z^2+\frac{1}{3!}z^3+\cdots=e^z$ 之收敛半径(∞)指在复数域内来说的, 故 $B(t)$ 之特征值与一幂级数之收敛半径无从比较, 因二者不属同样性质的量, 因之按照上述定义 e^A 之方法, 固可仍定义 e^{tA}, 但当 $B(t)$ 为 t 之 n 级连续矩阵时就不便定义 $e^{B(t)}$ 了, 由是 $e^{\int A(t)dt}$ 之意义在上述之定义下就不明确了, 也就是说按照原来的方法无从定义 $e^{\int A(t)dt}$. 所以我们不得不采取另一种思想方法.

由于 (21) 中第二个式子的启发, 我们会想到去找一个 n 级连续矩阵 $F(t)$, 使(19)之解向量为

$$X=F(t)^{-1}\cdot\int F(t)B(t)dt.$$

于是就要看(21)中 $f(t)$ 具备哪些特征性质, 使我们今后的思路有所遵循. 因 $f(t)=e^{-\int a(t)dt}$, 故

i) $\frac{d}{dt}f(t)=-f(t)\cdot a(t)[=-a(t)\cdot f(t)]$,

ii) $f(t)\neq 0$, 对 t 可能取的一切值.

这性质 i)、ii) 对于解(19)有很大的启示, 因为据此, 我们就先承认有一个 n 级矩阵 $F(t)$, 对 t 之所有值满足 $\dot{F}(t)=$

$-\boldsymbol{F}(t)\cdot\boldsymbol{A}(t)$ 且 $|\boldsymbol{F}(t)|\neq 0$，然后根据这样一个 $\boldsymbol{F}(t)$ 的存在性就能够推导 (19) 之解的存在与唯一，推演的步骤与求(20)之解的演算步骤大致相同．详述于下：

先假定存在一个 n 级连续矩阵 $\boldsymbol{F}(t)$，使对所有 t 值都满足

$$\dot{\boldsymbol{F}}(t)=-\boldsymbol{F}(t)\cdot\boldsymbol{A}(t) \tag{22}$$

与行列式

$$|\boldsymbol{F}(t)|\neq 0. \tag{23}$$

如果(19)有一个解向量 $\boldsymbol{X}$，则由(22)式有

$$\begin{aligned}\frac{d}{dt}\{\boldsymbol{F}(t)\boldsymbol{X}\}&=\dot{\boldsymbol{F}}(t)\cdot\boldsymbol{X}+\boldsymbol{F}(t)\cdot\dot{\boldsymbol{X}}\\&=-\boldsymbol{F}(t)\cdot\boldsymbol{A}(t)\cdot\boldsymbol{X}+\boldsymbol{F}(t)\cdot\{\boldsymbol{A}(t)\cdot\boldsymbol{X}+\boldsymbol{B}(t)\}\\&=\boldsymbol{F}(t)\cdot\boldsymbol{B}(t),\end{aligned}$$

故 $$\boldsymbol{F}(t)\boldsymbol{X}=\int_0^t\boldsymbol{F}(u)\boldsymbol{B}(u)du+\boldsymbol{F}(0)\cdot\boldsymbol{X}(0),$$

式中 $\boldsymbol{X}(0)=\begin{pmatrix}c_1\\\vdots\\c_n\end{pmatrix}$．因之，由(23)式可得

$$\boldsymbol{X}=\boldsymbol{F}(t)^{-1}\cdot\left\{\int_0^t\boldsymbol{F}(u)\boldsymbol{B}(u)du+\boldsymbol{F}(0)\cdot\boldsymbol{X}(0)\right\}. \tag{24}$$

这说明了：(19)如有解，则解是唯一的，且为(24)．

反之，易验证(24)又确是(19)的解．事实上，从(24)即得

$$\boldsymbol{F}(t)\cdot\boldsymbol{X}=\int_0^t\boldsymbol{F}(u)\boldsymbol{B}(u)du+\boldsymbol{F}(0)\cdot\boldsymbol{X}(0),$$

求导数，就有

$$\dot{\boldsymbol{F}}(t)\cdot\boldsymbol{X}+\boldsymbol{F}(t)\cdot\dot{\boldsymbol{X}}=\boldsymbol{F}(t)\cdot\boldsymbol{B}(t),$$

再以(22)式代入之，则得

$$-\boldsymbol{F}(t)\cdot\boldsymbol{A}(t)\cdot\boldsymbol{X}+\boldsymbol{F}(t)\cdot\dot{\boldsymbol{X}}=\boldsymbol{F}(t)\cdot\boldsymbol{B}(t);$$

由(23)式知 $\boldsymbol{F}(t)^{-1}$ 存在，故以 $\boldsymbol{F}(t)^{-1}$ 左乘后就得到

$$-\boldsymbol{A}(t)\cdot\boldsymbol{X}+\dot{\boldsymbol{X}}=\boldsymbol{B}(t),$$

即 $\dot{\boldsymbol{X}}=\boldsymbol{A}(t)\cdot\boldsymbol{X}+\boldsymbol{B}(t)$，这就是(19)式。

读者可能还会提出另一个想法，那就是当适合(22)与(23)之 $\boldsymbol{F}(t)$ 存在时，则(19)之解存在且唯一，而为(24)；然(22)之思想根源在 $f(t)$ 之性质 i)，而 $f(t)$ 之性质 i) 固可为 $\frac{d}{dt}f(t)=-f(t)\cdot a(t)$，但也可以为 $\frac{d}{dt}f(t)=-a(t)\cdot f(t)$；(22)式只不过是由前一表示法所启示，若从后一表示法之启示而使 $\dot{\boldsymbol{F}}(t)=-\boldsymbol{A}(t)\cdot\boldsymbol{F}(t)$，又怎样呢？这时，就应有

$$\begin{aligned}\frac{d}{dt}\{\boldsymbol{F}(t)\cdot\boldsymbol{X}\}&=\dot{\boldsymbol{F}}(t)\cdot\boldsymbol{X}+\boldsymbol{F}(t)\cdot\dot{\boldsymbol{X}}\\&=-\boldsymbol{A}(t)\boldsymbol{F}(t)\boldsymbol{X}+\boldsymbol{F}(t)\cdot\{\boldsymbol{A}(t)\boldsymbol{X}+\boldsymbol{B}(t)\},\end{aligned}$$

故结果未必一定为 $\boldsymbol{F}(t)\boldsymbol{B}(t)$，因之也就推不出(24)式。这说明了后一种思路失败了，得不到预期的结果。

注意：(24)显然是(21)之推广。

于是，剩下来要解决的就是要证明满足(22)与(23)之 $\boldsymbol{F}(t)$ 的存在性。

首先来谈求 $\boldsymbol{F}(t)$ 之思想方法的根源。当然还是要从 $f(t)=e^{-\int a(t)dt}$ 之性质来谈起，便于发现问题。$f(t)$ 除了具备上述的 i)，ii) 两个特性外，如果还要研究 $f(t)$ 之性质，就要考虑到分析学中常用的方法，即展开为无穷级数的方法。然由指数函数之展开式则知可以形式地把 $f(t)$ 写为

$$\begin{aligned}f(t)&=1-\int a(t)dt+\frac{1}{2!}\left[\int a(t)dt\right]^2-\cdots\\&=1-\int_0^t a(u)du+\frac{1}{2!}\int_0^t\left[\int_0^t a(u)du\right]a(u)du-\cdots,\end{aligned}$$

于是我们也想把需要求得的矩阵 $\boldsymbol{F}(t)$ 当做一个收敛矩阵之

和，因此根据上面的$f(t)$之表示方法，得到了下面的求作矩阵级数之和的思路.

用递归公式:

$$\boldsymbol{F}_0(t)=\boldsymbol{E},$$

$$\boldsymbol{F}_1(t)=\int_0^t \boldsymbol{F}_0(u)\boldsymbol{A}(u)du,$$

$$\boldsymbol{F}_2(t)=\int_0^t \boldsymbol{F}_1(u)\boldsymbol{A}(u)du,\ \cdots,$$

即 $$\boldsymbol{F}_{k+1}(t)=\int_0^t \boldsymbol{F}_k(u)\boldsymbol{A}(u)du \quad (k=0,\ 1,\ 2,\ \cdots),$$

而定义$\boldsymbol{F}_0(t)$, $\boldsymbol{F}_1(t)$, $\boldsymbol{F}_2(t)$, $\cdots$. 因$\boldsymbol{A}(t)$与$\boldsymbol{F}_0(t)=\boldsymbol{E}$均连续，故$\boldsymbol{F}_{k+1}(t)$也都是连续的$(k=0,\ 1,\ 2,\ \cdots)$，于是有了连续矩阵的一个序列$\{\boldsymbol{F}_k(t)\}$. 每个$\boldsymbol{F}_i$不仅仅是连续的，且实际上还是可微的，而有

$$\dot{\boldsymbol{F}}_{k+1}(t)=\boldsymbol{F}_k(t)\cdot\boldsymbol{A}(t),\ (k=0,\ 1,\ 2,\ \cdots).$$

再作矩阵级数

$$\boldsymbol{F}_0(t)-\boldsymbol{F}_1(t)+\boldsymbol{F}_2(t)-\boldsymbol{F}_3(t)+\cdots \tag{25}$$

时，我们希望级数(25)收敛且有和$\boldsymbol{F}(t)$并适合(22)与(23)式. 为此，首先解决的是(25)必需为一致收敛的(因为要求$\boldsymbol{F}(t)$是连续的).

用归纳法，可以证明

$$\|\boldsymbol{F}_k(t)\|\leqslant\frac{1}{k!}\cdot\{n\cdot|t|\cdot\overline{bd}\|\boldsymbol{A}(t)\|\}^k,$$

$$(k=0,\ 1,\ 2,\ \cdots), \tag{26}$$

式中$\overline{bd}\|\boldsymbol{A}(t)\|$表示$\|\boldsymbol{A}(t)\|$之上界——因$\boldsymbol{A}(t)$连续，故$\|\boldsymbol{A}(t)\|$之上界存在.

事实上，$k=0$时，(26)显然正确[这里要先假定$0^0=1$]. 今假定对某$k\geqslant 0$，(26)式为真. 于是，由分析学的知识，有

$$\|F_{k+1}(t)\|=\left\|\int_0^t F_k(u)A(u)du\right\|\leqslant\left|\int_0^t \|F_k(u)\cdot A(u)\|du\right|^{\dagger};$$

但 $\|F_k(u)A(u)\|=\max\limits_{1\leqslant i,j\leqslant n}\left|\sum\limits_{s=1}^{n}(F_k(u))_{is}(A(u))_{sj}\right|$

$$\leqslant n\cdot\|F_k(u)\|\cdot\|A(u)\|\leqslant n\cdot\|F_k(u)\|\cdot\overline{bd}\|A(u)\|,$$

故代入上式中，就有

$$\|F_{k+1}(t)\|\leqslant\left|\int_0^t n\cdot\|F_k(u)\|\cdot\overline{bd}\|A(u)\|\cdot du\right|,$$

因之，据归纳法之假设，以(26)代入之，即得

$$\begin{aligned}\|F_{k+1}(t)\|&\leqslant\left|\int_0^t n\cdot\frac{1}{k!}\{n\cdot|u|\cdot\overline{bd}\|A(u)\|\}^k\cdot\overline{bd}\|A(u)\|\cdot du\right|\\&=\frac{1}{k!}\cdot\{n\cdot\overline{bd}\|A(u)\|\}^{k+1}\cdot\left|\int_0^t|u|^k du\right|\\&=\frac{1}{k!}\cdot\{n\cdot\overline{bd}\|A(u)\|\}^{k+1}\cdot\frac{|t|^{k+1}}{k+1}\\&=\frac{1}{(k+1)!}\{n\cdot|t|\cdot\overline{bd}\|A(u)\|\}^{k+1},\end{aligned}$$

这就用归纳法证明了(26)。

于是，若令 $c_k=\frac{1}{k!}\{n\cdot\beta\cdot\overline{bd}\|A(t)\|\}^k$，则正数 c_k 与 t 无关，且由(26)式显然可知 $\|F_k(t)\|\leqslant c_k$ $(k=0,1,2,\cdots)$。但 $\sum\limits_{k=0}^{\infty}c_k$ 为指数级数，故必收敛，因之据级数理论可知(25)为一致收敛的。

---

† $F_k(u)A(u)$ 之连续 $\Rightarrow(F_k(u)\cdot A(u))_{ij}$ 连续 $\Rightarrow|(F_k(u)A(u))_{ij}|$ 连续 $\Rightarrow$ $\max\limits_{1\leqslant i,j\leqslant n}|(F_k(u)A(u))_{ij}|$ 连续 $\Rightarrow\|F_k(u)A(u)\|$ 连续，故 $\int_0^t\|F_k(u)A(u)\|du$ 存在。又 $(F_k(u)A(u))_{ij}\leqslant\|F_k(u)A(u)\|\Rightarrow\left|\int_0^t(F_k(u)A(u))_{ij}du\right|\leqslant$ $\left|\int_0^t\|F_k(u)A(u)\|du\right|$，因而

$$\left\|\int_0^t F_k(u)A(u)du\right\|=\max_{1\leqslant i,j\leqslant n}\left|\int_0^t(F_k(u)A(u))_{ij}du\right|\leqslant\left|\int_0^t\|F_k(u)A(u)\|du\right|.$$

既已证明(25)之一致收敛性，故其和 $\boldsymbol{F}(t)$ 为连续函数且可逐项微分，于是利用关系式 $\dot{\boldsymbol{F}}_{k+1}(t)=\boldsymbol{F}_k(t)\cdot\boldsymbol{A}(t)$ 可知

$$\begin{aligned}\dot{\boldsymbol{F}}(t)&=-\boldsymbol{F}_0(t)\cdot\boldsymbol{A}(t)+\boldsymbol{F}_1(t)\cdot\boldsymbol{A}(t)\\&\quad-\boldsymbol{F}_2(t)\cdot\boldsymbol{A}(t)+\boldsymbol{F}_3(t)\cdot\boldsymbol{A}(t)-\cdots\\&=-\{\boldsymbol{F}_0(t)-\boldsymbol{F}_1(t)+\boldsymbol{F}_2(t)-\boldsymbol{F}_3(t)\\&\quad+\cdots\}\cdot\boldsymbol{A}(t)=-\boldsymbol{F}(t)\cdot\boldsymbol{A}(t)\end{aligned}$$

(参见第 2 节的性质 3°)，证明了这样的 $\boldsymbol{F}(t)$ 确满足(22)式。因之，剩下的问题是要证明这个 $\boldsymbol{F}(t)$ 还需满足(23)式。

为此，用递归公式

$$\boldsymbol{G}_0(t)=\boldsymbol{E},$$

$$\boldsymbol{G}_{k+1}(t)=\int_0^t\boldsymbol{A}(u)\boldsymbol{G}_k(u)du,\ (k=0,\ 1,\ 2,\ \cdots)$$

来定义 $\boldsymbol{G}_0(t)$, $\boldsymbol{G}_1(t)$, $\boldsymbol{G}_2(t)$, $\cdots$. 与上述关于 $\{\boldsymbol{F}_k(t)\}$ 之理论完全类似地，又可以推得序列 $\{\boldsymbol{G}_k(t)\}$ 之相应的一些结果，因之得知矩阵

$$\boldsymbol{G}(t)=\sum_{k=0}^{\infty}\boldsymbol{G}_k(t)$$

满足关系式 $\dot{\boldsymbol{G}}(t)=\boldsymbol{A}(t)\cdot\boldsymbol{G}(t)$. 于是得：

$$\begin{aligned}\frac{d}{dt}\{\boldsymbol{F}(t)\boldsymbol{G}(t)\}&=\dot{\boldsymbol{F}}(t)\boldsymbol{G}(t)+\boldsymbol{F}(t)\cdot\dot{\boldsymbol{G}}(t)\\&=-\boldsymbol{F}(t)\boldsymbol{A}(t)\boldsymbol{G}(t)+\boldsymbol{F}(t)\boldsymbol{A}(t)\boldsymbol{G}(t)=\boldsymbol{0},\end{aligned}$$

表示着 $\boldsymbol{F}(t)\boldsymbol{G}(t)$ 与 t 无关，因之不论 t 为何值常有

$$\boldsymbol{F}(t)\boldsymbol{G}(t)=\boldsymbol{F}(0)\boldsymbol{G}(0).$$

但 $\boldsymbol{F}(0)=\boldsymbol{E}=\boldsymbol{G}(0)$ 是显然的，故不论 t 为何值常有

$$\boldsymbol{F}(t)\boldsymbol{G}(t)=\boldsymbol{E},$$

这就证明了(23). 至此，问题完全被解决了. 所以证得了下面的：

定理 2 满足初始条件 $x_1(0)=c_1$, $\cdots$, $x_n(0)=c_n$ 的一阶

变系数的线性非齐次微分方程组

$$
\begin{cases}
\frac{dx_1}{dt}=a_{11}(t)x_1+\cdots+a_{1n}(t)x_n+b_1(t),\\
\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\
\frac{dx_n}{dt}=a_{n1}(t)x_1+\cdots+a_{nn}(t)x_n+b_n(t),
\end{cases}
\tag{27}
$$

有且只有唯一的一组函数 $x_1(t)$, …, $x_n(t)$ 为其解，式中 $a_{ij}(t)$, $b_i(t)$ 都是 t 之连续函数(i, $j=1, 2, \cdots, n$)，而 c_1, c_2, …, c_n 是任意常数.

由定理 2 还可直接推得定理 1. 事实上，这时 $\boldsymbol{A}(t)=\boldsymbol{A}=(a_{ij})$ 为常量矩阵，故 $\boldsymbol{F}_0(t)=\boldsymbol{E}$, $\boldsymbol{F}_{k+1}(t)=\left[\int_0^t \boldsymbol{F}_k(u)du\right]\boldsymbol{A}$, 由是 $\boldsymbol{F}_1(t)=t\boldsymbol{A}$, $\boldsymbol{F}_2(t)=\frac{t^2\boldsymbol{A}^2}{2!}$, $\boldsymbol{F}_3(t)=\frac{t^3\boldsymbol{A}^3}{3!}$, …, $\boldsymbol{F}_k(t)=\frac{t^k\boldsymbol{A}^k}{k!}$, …, 故 $\boldsymbol{F}(t)=\boldsymbol{F}_0(t)-\boldsymbol{F}_1(t)+\boldsymbol{F}_2(t)-\cdots=e^{-t\boldsymbol{A}}$. 这时，又因为有 $\boldsymbol{B}(t)=\begin{pmatrix}0\\0\\\vdots\\0\end{pmatrix}$, 故 $\int_0^t \boldsymbol{F}(u)\boldsymbol{B}(u)du=\begin{pmatrix}0\\0\\\vdots\\0\end{pmatrix}$, 于是由 (24) 即知这时 (27) 的唯一解是 $\boldsymbol{X}=\boldsymbol{F}(t)^{-1}\boldsymbol{F}(0)\boldsymbol{X}(0)=e^{t\boldsymbol{A}}\cdot\boldsymbol{X}(0)$, 此即定理 1.

第 5 节　迭代法收敛的条件

我们知道，线性方程组

$$
\begin{cases}
c_{11}x_1+\cdots+c_{1n}x_n=b_1,\\
\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\
c_{n1}x_1+\cdots+c_{nn}x_n=b_n,
\end{cases}
\tag{28}
$$

根据克莱姆定理，当行列式 $\begin{vmatrix} c_{11} & \cdots & c_{1n} \\ \vdots & & \vdots \\ c_{n1} & \cdots & c_{nn} \end{vmatrix} \neq 0$ 时，有唯一组解。但在实际工作中，一般不用克莱姆定理，原因是计算工作量太大，而通常采用消去法。消去法的优点是：能有限多次地可得到答案；解答具精确性。但缺点是消去法的程序复杂，即在计算机上的存贮量较大。一个实际问题为要在计算机上进行计算，就得尽可能地使程序简单，即是使存贮量较小。为此，通常是采用迭代法，即将(28)式改写为等价形式：

$$\begin{cases} x_1 = a_{11}x_1 + \cdots + a_{1n}x_n + b_1, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ x_n = a_{n1}x_1 + \cdots + a_{nn}x_n + b_n, \end{cases}$$

或写为矩阵之表达形式

$$\boldsymbol{X} = \boldsymbol{A}\boldsymbol{X} + \boldsymbol{B}, \tag{29}$$

式中 $\boldsymbol{X} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$, $\boldsymbol{B} = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$, $\boldsymbol{A} = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}$.

首先，取$(x_1, \cdots, x_n)$之任一组初始值$(x_1^{(0)}, x_2^{(0)}, \cdots, x_n^{(0)})$代入(29)之右端，而令左端之值为$(x_1^{(1)}, x_2^{(1)}, \cdots, x_n^{(1)})$，叫做第一次迭代值。再以这组第一次迭代值$(x_1^{(1)}, x_2^{(1)}, \cdots, x_n^{(1)})$代入(29)之右端，而令左端相应之值为$(x_1^{(2)}, x_2^{(2)}, \cdots, x_n^{(2)})$，叫做第二次迭代值。继续这样作下去，可得第$k$次迭代值$(x_1^{(k)}, x_2^{(k)}, \cdots, x_n^{(k)})$。如令 $\boldsymbol{X}^{(k)} = \begin{pmatrix} x_1^{(k)} \\ x_2^{(k)} \\ \vdots \\ x_n^{(k)} \end{pmatrix}$，则 $\boldsymbol{X}^{(k+1)} = \boldsymbol{A}\boldsymbol{X}^{(k)} +$

$\boldsymbol{B}$. 若 $\lim\limits_{k\to\infty}\boldsymbol{X}^{(k)}=\boldsymbol{Y}=\begin{pmatrix}y_1\\ y_2\\ \vdots\\ y_n\end{pmatrix}$，即 $\lim\limits_{k\to\infty}x_i^{(k)}=y_i(i=1, 2, \cdots, n)$，则 $\boldsymbol{Y}$ 为(29)之解.

这是因为：从 $\boldsymbol{X}^{(k+1)}=\boldsymbol{A}\boldsymbol{X}^{(k)}+\boldsymbol{B}$，两端取极限，得 $\boldsymbol{Y}=\lim\limits_{k\to\infty}\boldsymbol{X}^{(k+1)}=\boldsymbol{A}\cdot\lim\limits_{k\to\infty}\boldsymbol{X}^{(k)}+\boldsymbol{B}=\boldsymbol{A}\boldsymbol{Y}+\boldsymbol{B}$.

于是，接着要解决的就是极限 $\lim\limits_{k\to\infty}\boldsymbol{X}^{(k)}$ 存在的条件．如果极限存在，则即求出(29)即(28)之解，且这时程序也简单，而初始值 $(x_1^{(0)}, x_2^{(0)}, \cdots, x_n^{(0)})$ 又可任意．这是迭代法的优点．当极限 $\lim\limits_{k\to\infty}\boldsymbol{X}^{(k)}$ 存在时，看实际的需要程度，我们求到第 k 次迭代值 $(x_1^{(k)}, \cdots, x_n^{(k)})$ 就作为(29)的解(对某一个适当的 k)；虽然这个第 k 次迭代值 $(x_1^{(k)}, \cdots, x_n^{(k)})$ 不是(29)的精确解，但随着 k 之增大，$\boldsymbol{X}^{(k)}$ 与精确解之误差也就变小了，故已可解决实际问题．因而，下面要解决的问题是何时才有 $\lim\limits_{k\to\infty}\boldsymbol{X}^{(k)}$ 存在呢？

因

$$\boldsymbol{A}\boldsymbol{X}^{(0)}+\boldsymbol{B}=\boldsymbol{X}^{(1)},$$

$$\begin{aligned}\boldsymbol{A}^2\boldsymbol{X}^{(0)}+\boldsymbol{A}\boldsymbol{B}+\boldsymbol{B}&=\boldsymbol{A}(\boldsymbol{A}\boldsymbol{X}^{(0)}+\boldsymbol{B})+\boldsymbol{B}\\&=\boldsymbol{A}\boldsymbol{X}^{(1)}+\boldsymbol{B}=\boldsymbol{X}^{(2)},\end{aligned}$$

而由归纳法可知

$$\begin{aligned}\boldsymbol{X}^{(k)}&=\boldsymbol{A}^k\boldsymbol{X}^{(0)}+(\boldsymbol{A}^{k-1}+\boldsymbol{A}^{k-2}+\cdots+\boldsymbol{A}^2+\boldsymbol{A}+\boldsymbol{E})\boldsymbol{B}\\&=\boldsymbol{A}^k\boldsymbol{X}^{(0)}+\boldsymbol{S}_k\boldsymbol{B},\end{aligned}$$

但

$$\boldsymbol{S}_k=\boldsymbol{A}^{k-1}+\boldsymbol{A}^{k-2}+\cdots+\boldsymbol{A}^2+\boldsymbol{A}+\boldsymbol{E}.$$

于是，因 $\boldsymbol{X}^{(0)}$ 与 $\boldsymbol{B}$ 与 k 无关，故当 $\lim\limits_{k\to\infty}\boldsymbol{A}^k$ 与 $\lim\limits_{k\to\infty}\boldsymbol{S}_k$ 存在时，$\lim\limits_{k\to\infty}\boldsymbol{X}^{(k)}$ 当然存在. 然 $\lim\limits_{k\to\infty}\boldsymbol{S}_k$ 之存在即说明了级数

$$\sum_{m=0}^{\infty}\boldsymbol{A}^m=\boldsymbol{E}+\boldsymbol{A}+\boldsymbol{A}^2+\boldsymbol{A}^3+\cdots$$

是收敛的. 我们已经知道：当 $\boldsymbol{A}$ 之特征值的模都小于 1 时，$\sum\limits_{m=0}^{\infty}\boldsymbol{A}^m$ 收敛(第 2 节的定理之推论 2) 且 $\lim\limits_{k\to\infty}\boldsymbol{S}_k=(\boldsymbol{E}-\boldsymbol{A})^{-1}$，又这时 $\lim\limits_{k\to\infty}\boldsymbol{A}^k$ 存在且为 $\mathbf{0}$(第 1 节的定理 2). 于是，若 $\boldsymbol{A}$ 之特征值全在复平面之单位圆内时，则 $\lim\limits_{k\to\infty}\boldsymbol{X}^{(k)}$ 存在且等于 $(\boldsymbol{E}-\boldsymbol{A})^{-1}\cdot\boldsymbol{B}$，即 $\lim\limits_{k\to\infty}\boldsymbol{X}^{(k)}=(\boldsymbol{E}-\boldsymbol{A})^{-1}\cdot\boldsymbol{B}$.

反之，若 $\lim\limits_{k\to\infty}\boldsymbol{X}^{(k)}$ 存在，则因初始值 $\boldsymbol{X}^{(0)}$ 可任取，故特取 $\boldsymbol{X}^{(0)}=\mathbf{0}$ 时，则有 $\boldsymbol{X}^{(k)}=\boldsymbol{A}^k\boldsymbol{X}^{(0)}+\boldsymbol{S}_k\boldsymbol{B}=\boldsymbol{S}_k\boldsymbol{B}$，故从 $\lim\limits_{k\to\infty}\boldsymbol{X}^{(k)}$ 之存在得 $\lim\limits_{k\to\infty}\boldsymbol{S}_k$ 存在，即

$$\sum_{m=0}^{\infty}\boldsymbol{A}^m$$

是收敛的，因而据第 2 节之定理的推论 3，得知 $\boldsymbol{A}$ 之特征值全在复平面之单位圆内.

于是，证得了：

定理 1 当且仅当矩阵 $\boldsymbol{A}$ 之特征值的模都小于 1 时，方程组(28)之迭代形式(29)不论初始值 $\boldsymbol{X}^{(0)}$ 如何选取恒为收敛的，即 $\lim\limits_{k\to\infty}\boldsymbol{X}^{(k)}$ 存在，且等于 $(\boldsymbol{B}-\boldsymbol{A})^{-1}\cdot\boldsymbol{B}$.

定理 1 就是线代数计算方法中的所谓简单迭代法. 除了简单迭代法，还有所谓逐步迭代法，它的意义是这样的，即以任意初始值 $(x_1^{(0)}, x_2^{(0)}, \cdots, x_n^{(0)})$ 代入(29)之第一个方程的右边而令左端之值为 $x_1^{(1)}$，再以 $(x_1^{(1)}, x_2^{(0)}, \cdots, x_n^{(0)})$ 代入(29)之

第二个方程的右端而令左端之值为 $x_2^{(1)}$，一般若 $x_k^{(1)}$ 求得后，再以 $(x_1^{(1)}, \cdots, x_k^{(1)}, x_{k+1}^{(0)}, \cdots, x_n^{(0)})$ 代入(29)之第 $k+1$ 个方程的右端而令左端之值为 $x_{k+1}^{(1)}$，继续下去，直至求得 $x_{n-1}^{(1)}$ 并以 $(x_1^{(1)}, \cdots, x_{n-1}^{(1)}, x_n^{(0)})$ 代入(29)之最后一个方程的右端而令左端之值为 $x_n^{(1)}$. 象这样得到的 $(x_1^{(1)}, x_2^{(1)}, \cdots, x_n^{(1)})$ 叫做第一次逐步迭代值. 如果第 k 次逐步迭代值 $(x_1^{(k)}, x_2^{(k)}, \cdots, x_n^{(k)})$ 已求得，再以之代入(29)之第一个方程的右端，而令左端之值为 $x_1^{(k+1)}$，然后用 $x_1^{(k+1)}$ 去替换 $x_1^{(k)}$ 而以 $(x_1^{(k+1)}, x_2^{(k)}, \cdots, x_n^{(k)})$ 代入(29)之第二个方程的右端并令左端之值为 $x_2^{(k+1)}$，继续下去，直至求得了 $x_{n-1}^{(k+1)}$，再以 $(x_1^{(k+1)}, \cdots, x_{n-1}^{(k+1)}, x_n^{(k)})$ 代入(29)之末式的右端而令左端之值为 $x_n^{(k+1)}$，这样就又求得了第 $k+1$ 次逐步迭代值 $(x_1^{(k+1)}, x_2^{(k+1)}, \cdots, x_n^{(k+1)})$. 象这样的迭代方法叫做逐步迭代法. 采用逐步迭代法，可望收敛的速度快些，即希望用较小的 k 而使第 k 次逐步迭代值 $(x_1^{(k)}, x_2^{(k)}, \cdots, x_n^{(k)})$ 更接近于精确解 $(x_1^*, x_2^*, \cdots, x_n^*)$（若存在的话）.

今将上述逐步迭代法之步骤归纳于下：

$$\begin{cases} x_1^{(k+1)} = a_{11}x_1^{(k)} + a_{12}x_2^{(k)} + a_{13}x_3^{(k)} + \cdots \\ \qquad\qquad + a_{1,n-1}x_{n-1}^{(k)} + a_{1n}x_n^{(k)} + b_1, \\ x_2^{(k+1)} = a_{21}x_1^{(k+1)} + a_{22}x_2^{(k)} + a_{23}x_3^{(k)} + \cdots \\ \qquad\qquad + a_{2,n-1}x_{n-1}^{(k)} + a_{2n}x_n^{(k)} + b_2, \\ x_3^{(k+1)} = a_{31}x_1^{(k+1)} + a_{32}x_2^{(k+1)} + a_{33}x_3^{(k)} + \cdots \\ \qquad\qquad + a_{3,n-1}x_{n-1}^{(k)} + a_{3n}x_n^{(k)} + b_3, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ x_n^{(k+1)} = a_{n1}x_1^{(k+1)} + a_{n2}x_2^{(k+1)} + a_{n3}x_3^{(k+1)} + \cdots \\ \qquad\qquad + a_{n,n-1}x_{n-1}^{(k+1)} + a_{nn}x_n^{(k)} + b_n, \end{cases}$$

改写为矩阵之形式，为：

$$\begin{pmatrix} x_1^{(k+1)} \\ x_2^{(k+1)}-a_{21}x_1^{(k+1)} \\ x_3^{(k+1)}-a_{31}x_1^{(k+1)}-a_{32}x_2^{(k+1)} \\ \vdots \\ x_n^{(k+1)}-a_{n1}x_1^{(k+1)}-\cdots-a_{n,n-1}x_{n-1}^{(k+1)} \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1,n-1} & a_{1n} \\ 0 & a_{22} & a_{23} & \cdots & a_{2,n-1} & a_{2n} \\ 0 & 0 & a_{33} & \cdots & a_{3,n-1} & a_{3n} \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & a_{n-1,n-1} & a_{n-1,n} \\ 0 & 0 & 0 & \cdots & 0 & a_{nn} \end{pmatrix} \begin{pmatrix} x_1^{(k)} \\ x_2^{(k)} \\ x_3^{(k)} \\ \vdots \\ x_{n-1}^{(k)} \\ x_n^{(k)} \end{pmatrix} + \begin{pmatrix} b_1 \\ b_2 \\ b_3 \\ \vdots \\ b_{n-1} \\ b_n \end{pmatrix},$$

亦即

$$\begin{pmatrix} 1 & 0 & 0 & \cdots & 0 & 0 \\ -a_{21} & 1 & 0 & \cdots & 0 & 0 \\ -a_{31} & -a_{32} & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ -a_{n-1,1} & -a_{n-1,2} & -a_{n-1,3} & \cdots & 1 & 0 \\ -a_{n1} & -a_{n2} & -a_{n3} & \cdots & -a_{n,n-1} & 1 \end{pmatrix} \begin{pmatrix} x_1^{(k+1)} \\ x_2^{(k+1)} \\ x_3^{(k+1)} \\ \vdots \\ x_{n-1}^{(k+1)} \\ x_n^{(k+1)} \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1,n-1} & a_{1n} \\ 0 & a_{22} & a_{23} & \cdots & a_{2,n-1} & a_{2n} \\ 0 & 0 & a_{33} & \cdots & a_{3,n-1} & a_{3n} \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & a_{n-1,n-1} & a_{n-1,n} \\ 0 & 0 & 0 & \cdots & 0 & a_{nn} \end{pmatrix} \begin{pmatrix} x_1^{(k)} \\ x_2^{(k)} \\ x_3^{(k)} \\ \vdots \\ x_{n-1}^{(k)} \\ x_n^{(k)} \end{pmatrix} + \begin{pmatrix} b_1 \\ b_2 \\ b_3 \\ \vdots \\ b_{n-1} \\ b_n \end{pmatrix}$$

或
$$(\boldsymbol{E}-\boldsymbol{L})\boldsymbol{X}^{(k+1)}=\boldsymbol{V}\boldsymbol{X}^{(k)}+\boldsymbol{B},$$

式中
$$\boldsymbol{B}=\begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}, \quad \boldsymbol{X}^{(k)}=\begin{pmatrix} x_1^{(k)} \\ x_2^{(k)} \\ \vdots \\ x_n^{(k)} \end{pmatrix},$$

$$\boldsymbol{V}=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{pmatrix}\text{为上三角阵，}\boldsymbol{L}=\begin{pmatrix} 0 & 0 & 0 & \cdots & 0 \\ a_{21} & 0 & 0 & \cdots & 0 \\ a_{31} & a_{32} & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & a_{n3} & \cdots & 0 \end{pmatrix}$$

为下三角阵．显然，$\boldsymbol{E}-\boldsymbol{L}$ 为满秩，故又可将上式写为
$$\boldsymbol{X}^{(k+1)}=(\boldsymbol{E}-\boldsymbol{L})^{-1}\boldsymbol{V}\boldsymbol{X}^{(k)}+\boldsymbol{F}, \tag{30}$$
式中 $\boldsymbol{F}=(\boldsymbol{E}-\boldsymbol{L})^{-1}\boldsymbol{B}$.

这说明了逐步迭代法的形式已转化为简单迭代法之形(30)．当然，逐步迭代法如收敛，即 $\lim\limits_{k\to\infty}\boldsymbol{X}^{(k)}$ 存在(如令为 $\lim\limits_{k\to\infty}\boldsymbol{X}^{(k)}=\boldsymbol{X}^*$)，则当注意了 $\boldsymbol{L}+\boldsymbol{V}=\boldsymbol{A}$ 后从(30)式就有
$$\begin{aligned}\boldsymbol{X}^*&=\lim_{k\to\infty}\boldsymbol{X}^{(k+1)}\\&=(\boldsymbol{E}-\boldsymbol{L})^{-1}\boldsymbol{V}\cdot\lim_{k\to\infty}\boldsymbol{X}^{(k)}+\boldsymbol{F}\\&=(\boldsymbol{E}-\boldsymbol{L})^{-1}\cdot\boldsymbol{V}\cdot\boldsymbol{X}^*+\boldsymbol{F},\end{aligned}$$
即
$$(\boldsymbol{E}-\boldsymbol{L})\boldsymbol{X}^*=\boldsymbol{V}\boldsymbol{X}^*+(\boldsymbol{E}-\boldsymbol{L})\boldsymbol{F}=\boldsymbol{V}\boldsymbol{X}^*+\boldsymbol{B},$$
或
$$\boldsymbol{X}^*=(\boldsymbol{L}+\boldsymbol{V})\boldsymbol{X}^*+\boldsymbol{B}=\boldsymbol{A}\boldsymbol{X}^*+\boldsymbol{B},$$
即说明了 $\boldsymbol{X}^*$ 确为(29)之解．于是，问题是判定逐步迭代法的收敛条件，也就是判定(30)之收敛条件．然由定理 1，知所求条件为

$$|\lambda \boldsymbol{E}-(\boldsymbol{E}-\boldsymbol{L})^{-1}\boldsymbol{V}|=0$$

之根全在单位圆内，亦即 $|\lambda(\boldsymbol{E}-\boldsymbol{L})-\boldsymbol{V}|=0$ 之根全在单位圆内．故又证得了：

定理 2 方程组(28)之迭代形式(29)不论初始值 $\boldsymbol{X}^{(0)}$ 如何选取，关于逐步迭代恒为收敛的充要条件是多项式

$$\begin{vmatrix} \lambda-a_{11} & -a_{12} & -a_{13} & \cdots & -a_{1n} \\ -a_{21}\lambda & \lambda-a_{22} & -a_{23} & \cdots & -a_{2n} \\ -a_{31}\lambda & -a_{32}\lambda & \lambda-a_{33} & \cdots & -a_{3n} \\ \vdots & \vdots & \vdots & & \vdots \\ -a_{n1}\lambda & -a_{n2}\lambda & -a_{n3}\lambda & \cdots & \lambda-a_{nn} \end{vmatrix}=0$$

的根全在复平面的单位圆内．

附注　某迭代式可能用简单迭代法收敛，但用逐步迭代法则不收敛；也有用逐步迭代法收敛的，而用简单迭代法却不收敛．例如：对于：

$$\begin{cases} x_1=\dfrac{1}{2}x_1-\dfrac{1}{4}x_2+1, \\ x_2=3x_1+\dfrac{1}{2}x_2+2. \end{cases}$$

由于 $\boldsymbol{A}=\begin{pmatrix} \dfrac{1}{2} & -\dfrac{1}{4} \\ 3 & \dfrac{1}{2} \end{pmatrix}$ 之特征值为 $\dfrac{1\pm\sqrt{3}\,i}{2}$，其模 $=1$，故用简单迭代法则不收敛．

但这时

$$\boldsymbol{L}=\begin{pmatrix} 0 & 0 \\ 3 & 0 \end{pmatrix},\quad \boldsymbol{V}=\begin{pmatrix} \dfrac{1}{2} & -\dfrac{1}{4} \\ 0 & \dfrac{1}{2} \end{pmatrix},$$

故 $$|\lambda(\boldsymbol{E}-\boldsymbol{L})-\boldsymbol{V}|=\begin{vmatrix}\lambda-\frac{1}{2} & \frac{1}{4}\\ -3\lambda & \lambda-\frac{1}{2}\end{vmatrix}=0$$

之二根为 $\frac{1\pm\sqrt{15}i}{8}$，它们的模都是 $\frac{1}{2}$，故用逐步迭代法是收敛的.

而对于 $$\begin{cases}x_1=-3x_1+x_2+k_1,\\ x_2=-\frac{385}{64}x_1+2x_2+k_2.\end{cases}$$

由于 $\boldsymbol{A}=\begin{pmatrix}-3 & 1\\ -\frac{385}{64} & 2\end{pmatrix}$ 之特征值为 $-\frac{1}{2}\pm\frac{\sqrt{15}}{8}$，其模为 $\left|\frac{-4+\sqrt{15}}{8}\right|<1$ 及 $\left|\frac{-4-\sqrt{15}}{8}\right|<1$，故用简单迭代法则收敛.

但 $$\boldsymbol{L}=\begin{pmatrix}0 & 0\\ -\frac{385}{64} & 0\end{pmatrix},\quad \boldsymbol{V}=\begin{pmatrix}-3 & 1\\ 0 & 2\end{pmatrix},$$

故 $$|\lambda(\boldsymbol{E}-\boldsymbol{L})-\boldsymbol{V}|=\begin{vmatrix}\lambda+3 & -1\\ \frac{385}{64}\lambda & \lambda-2\end{vmatrix}=0$$

之二根为 $-\frac{449}{128}\pm\frac{1}{128}\sqrt{299905}$，显然 $-\frac{449}{128}-\frac{1}{128}\sqrt{299905}$ 之模大于 1，故用逐步迭代法则不收敛.

通常，在线代数计算中，往往将所与的方程组(28)若能变成迭代形式(29)，使矩阵 $\boldsymbol{A}$ 之每行(或列)各元之模的和都小于 1，则不论初始值 $X^{(0)}$ 怎样选取，确为收敛的，这在一般线代数计算方法的教科书中均有，这里就不重复了．现在直接利用定理 1 来解决这个问题，也就是：

定理3 若n级矩阵$\boldsymbol{A}=(a_{ij})$中$|a_{i1}|+|a_{i2}|+\cdots+|a_{in}|<1$ $(i=1, 2, \cdots, n)$ [或 $|a_{1j}|+|a_{2j}|+\cdots+|a_{nj}|<1(j=1, 2, \cdots, n)$]，则$\boldsymbol{A}$之每特征值的模小于1.

证明 若λ为$\boldsymbol{A}$之任一特征值，则有一非零向量$\boldsymbol{X}=\begin{pmatrix}x_1\\ \vdots\\ x_n\end{pmatrix}$，使$\boldsymbol{AX}=\lambda\boldsymbol{X}$. 如令$|x_k|=\max(|x_1|, \cdots, |x_n|)$，则$|x_k|>0$，且有

$$\lambda x_k=a_{k1}x_1+a_{k2}x_2+\cdots+a_{kn}x_n,$$

所以：

$$\begin{aligned}|\lambda x_k| &\leqslant |a_{k1}|\cdot|x_1|+|a_{k2}|\cdot|x_2|+\cdots+|a_{kn}|\cdot|x_n|\\ &\leqslant (|a_{k1}|+|a_{k2}|+\cdots+|a_{kn}|)\cdot|x_k|,\end{aligned}$$

故有

$$|\lambda|\leqslant|a_{k1}|+|a_{k2}|+\cdots+|a_{kn}|<1.$$

同理，若$|a_{1j}|+|a_{2j}|+\cdots+|a_{nj}|<1(j=1, 2, \cdots, n)$，则存在非零行向量$\boldsymbol{X}=(x_1, x_2, \cdots, x_n)$使$\lambda\boldsymbol{X}=\boldsymbol{XA}$，故当$|x_k|=\max(|x_1|, \cdots, |x_n|)$时，则$|x_k|>0$且有$\lambda x_k=a_{1k}x_1+a_{2k}x_2+\cdots+a_{nk}x_n$，因而有

$$|\lambda|\,|x_k|\leqslant(|a_{1k}|+|a_{2k}|+\cdots+|a_{nk}|)\cdot|x_k|,$$

故$|\lambda|\leqslant|a_{1k}|+|a_{2k}|+\cdots+|a_{nk}|<1$. 证完.

第十章　若干问题

本章将讲几个特殊问题，如直积、复合矩阵、特征值的估计、矩阵之极分解、全矩阵环之基，等等．它们有的在矩阵的理论研究中非常重要，有的是在应用中显示了价值．

第1节　矩阵的直积

先从简单的例子谈起．设有二元向量 (x_1, x_2) 与三元向量 (y_1, y_2, y_3)，它们分别经过二级矩阵与三级矩阵

$$\boldsymbol{A}=\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \quad 与 \quad \boldsymbol{B}=\begin{pmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{pmatrix}$$

的变换变成向量 (x_1', x_2') 与 (y_1', y_2', y_3')，即

$$\begin{pmatrix} x_1' \\ x_2' \end{pmatrix}=\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix},$$

$$\begin{pmatrix} y_1' \\ y_2' \\ y_3' \end{pmatrix}=\begin{pmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{pmatrix}\begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}.$$

现在来考察以这两个向量的分量的乘积为分量的六元向量 $(x_1y_1, x_1y_2, x_1y_3, x_2y_1, x_2y_2, x_2y_3)$，经过怎样的线性变换可以变成六元向量 $(x_1'y_1', x_1'y_2', x_1'y_3', x_2'y_1', x_2'y_2', x_2'y_3')$ 呢？

由假设　$x_i'=a_{i1}x_1+a_{i2}x_2 \quad (i=1, 2),$

$$y_j'=b_{j1}y_1+b_{j2}y_2+b_{j3}y_3 \quad (j=1, 2, 3).$$

故

$$
\begin{aligned}
x'_i y'_j &= (a_{i1}x_1 + a_{i2}x_2)(b_{j1}y_1 + b_{j2}y_2 + b_{j3}y_3) \\
&= a_{i1}b_{j1}x_1y_1 + a_{i1}b_{j2}x_1y_2 + a_{i1}b_{j3}x_1y_3 + a_{i2}b_{j1}x_2y_1 \\
&\quad + a_{i2}b_{j2}x_2y_2 + a_{i2}b_{j3}x_2y_3.
\end{aligned}
$$

于是，所求之变换的矩阵为六级矩阵

$$
P=\left(\begin{array}{ccc:ccc}
a_{11}b_{11} & a_{11}b_{12} & a_{11}b_{13} & a_{12}b_{11} & a_{12}b_{12} & a_{12}b_{13} \\
a_{11}b_{21} & a_{11}b_{22} & a_{11}b_{23} & a_{12}b_{21} & a_{12}b_{22} & a_{12}b_{23} \\
a_{11}b_{31} & a_{11}b_{32} & a_{11}b_{33} & a_{12}b_{31} & a_{12}b_{32} & a_{12}b_{33} \\
\hdashline
a_{21}b_{11} & a_{21}b_{12} & a_{21}b_{13} & a_{22}b_{11} & a_{22}b_{12} & a_{22}b_{13} \\
a_{21}b_{21} & a_{21}b_{22} & a_{21}b_{23} & a_{22}b_{21} & a_{22}b_{22} & a_{22}b_{23} \\
a_{21}b_{31} & a_{21}b_{32} & a_{21}b_{33} & a_{22}b_{31} & a_{22}b_{32} & a_{22}b_{33}
\end{array}\right).
$$

这个六级矩阵可简写为

$$
P=\begin{pmatrix} a_{11}B & a_{12}B \\ a_{21}B & a_{22}B \end{pmatrix},
$$

叫做 A 与 B 之直积，用记号 $A\times B$ 来表示。

同样，变换 B 与 A 之直积

$$
Q=\begin{pmatrix} b_{11}A & b_{12}A & b_{13}A \\ b_{21}A & b_{22}A & b_{23}A \\ b_{31}A & b_{32}A & b_{33}A \end{pmatrix}=B\times A
$$

使六元向量 $(x_1y_1,\ x_2y_1,\ x_1y_2,\ x_2y_2,\ x_1y_3,\ x_2y_3)$ 变为 $(x'_1y'_1,\ x'_2y'_1,\ x'_1y'_2,\ x'_2y'_2,\ x'_1y'_3,\ x'_2y'_3)$。

一般，设

$$
A=\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} \quad 和 \quad B=\begin{pmatrix} b_{11} & \cdots & b_{1m} \\ \vdots & & \vdots \\ b_{m1} & \cdots & b_{mm} \end{pmatrix}
$$

分别为 n 级和 m 级的两个矩阵，那末

$$A\times B=\begin{pmatrix} a_{11}B & a_{12}B & \cdots & a_{1n}B \\ a_{21}B & a_{22}B & \cdots & a_{2n}B \\ \vdots & \vdots & & \vdots \\ a_{n1}B & a_{n2}B & \cdots & a_{nn}B \end{pmatrix}$$

和
$$B\times A=\begin{pmatrix} b_{11}A & b_{12}A & \cdots & b_{1m}A \\ b_{21}A & b_{22}A & \cdots & b_{2m}A \\ \vdots & \vdots & & \vdots \\ b_{m1}A & b_{m2}A & \cdots & b_{mm}A \end{pmatrix}$$

分别叫做 A 与 B 之直积和 B 与 A 之直积. 显然,

$$E_n\times E_m=E_{mn}=E_m\times E_n,$$

这里的 E_i 表示 i 级单位矩阵. $A\times B$ 与 $B\times A$ 虽然都是 mn 级矩阵, 但它们一般是不相等的, 即 $A\times B\neq B\times A$.

下面来叙述有关直积的一些基本性质.

设 A 和 B 分别为 n 级和 m 级矩阵, 令 $P=A\times B$. 考察 P 之转置矩阵 P'. 这时, 显然可以看出, 不仅仅是 a_{ij} 与 a_{ji} 的位置要互相调换, 而且 B 中元素所在行与列的位置也要互相调换, 这就是说

$$P'=\begin{pmatrix} a_{11}B' & a_{21}B' & \cdots & a_{n1}B' \\ a_{12}B' & a_{22}B' & \cdots & a_{n2}B' \\ \vdots & \vdots & & \vdots \\ a_{1n}B' & a_{2n}B' & \cdots & a_{nn}B' \end{pmatrix}=A'\times B'.$$

故得:

定理 1 设 $P=A\times B$, 则 $P'=A'\times B'$. 即 $(A\times B)'=A'\times B'$.

再设 $C=(c_{ij})$ 与 $D=(d_{ij})$ 也分别是 n 级与 m 级的矩阵, 并作

$$Q=C\times D=\begin{pmatrix} c_{11}D & c_{12}D & \cdots & c_{1n}D \\ c_{21}D & c_{22}D & \cdots & c_{2n}D \\ \vdots & \vdots & & \vdots \\ c_{n1}D & c_{n2}D & \cdots & c_{nn}D \end{pmatrix}.$$

于是 PQ 是 mn 级矩阵. 它的第 $(\lambda-1)m+i$ 行第 $(\mu-1)m+j$ 列交叉处的元素, 等于 P 之第 $(\lambda-1)m+i$ 行上诸元素 $a_{\lambda1}b_{i1}, \cdots, a_{\lambda1}b_{im}, a_{\lambda2}b_{i1}, \cdots, a_{\lambda2}b_{im}, \cdots, a_{\lambda n}b_{i1}, \cdots, a_{\lambda n}b_{im}$ 分别与 Q 之第 $(\mu-1)m+j$ 列上各对应元 $c_{1\mu}d_{1j}, \cdots, c_{1\mu}d_{mj}, c_{2\mu}d_{1j}, \cdots, c_{2\mu}d_{mj}, \cdots, c_{n\mu}d_{1j}, \cdots, c_{n\mu}d_{mj}$ 的乘积之和, 即为

$$\begin{aligned}&\{(a_{\lambda1}b_{i1})(c_{1\mu}d_{1j})+\cdots+(a_{\lambda1}b_{im})(c_{1\mu}d_{mj})\}+\cdots\\&\quad+\{(a_{\lambda n}b_{i1})(c_{n\mu}d_{1j})+\cdots+(a_{\lambda n}b_{im})(c_{n\mu}d_{mj})\}\\&=a_{\lambda1}c_{1\mu}f_{ij}+\cdots+a_{\lambda n}c_{n\mu}f_{ij}=e_{\lambda\mu}f_{ij},\end{aligned}$$

这里 $$f_{ij}=\sum_{t=1}^{m}b_{it}d_{tj},\qquad e_{\lambda\mu}=\sum_{s=1}^{n}a_{\lambda s}c_{s\mu},$$

即 $$AC=(e_{ij}),\quad BD=(f_{ij}).$$

因之,

$$PQ=\begin{pmatrix} e_{11}BD & e_{12}BD & \cdots & e_{1n}BD \\ e_{21}BD & e_{22}BD & \cdots & e_{2n}BD \\ \vdots & \vdots & & \vdots \\ e_{n1}BD & e_{n2}BD & \cdots & e_{nn}BD \end{pmatrix}=AC\times BD.$$

故又得:

定理 2 $(A\times B)\cdot(C\times D)=AC\times BD$, 但 A 与 C 为同级矩阵, B 与 D 亦为同级矩阵.

从直积之定义和定理 1 与定理 2, 容易推知下列各定理:

定理 3 设 A 与 B 都是对角矩阵, 则 $A\times B$ 也是对角矩阵.

定理 4 设 A 与 B 是同型的三角矩阵, 则 $A\times B$ 也是同型的三角矩阵. (同型是指同为上三角阵, 或同为下三

角阵.)

定理 5 $(\boldsymbol{A}\times\boldsymbol{B})\times\boldsymbol{C}=\boldsymbol{A}\times(\boldsymbol{B}\times\boldsymbol{C})$.

证明 由定义易知:

$$(\boldsymbol{A}\times\boldsymbol{B})\times\boldsymbol{C}=\begin{pmatrix}a_{11}\boldsymbol{B} & \cdots & a_{1n}\boldsymbol{B}\\ \vdots & & \vdots\\ a_{n1}\boldsymbol{B} & \cdots & a_{nn}\boldsymbol{B}\end{pmatrix}\times\boldsymbol{C}$$

$$=\begin{pmatrix}a_{11}(\boldsymbol{B}\times\boldsymbol{C}) & \cdots & a_{1n}(\boldsymbol{B}\times\boldsymbol{C})\\ \vdots & & \vdots\\ a_{n1}(\boldsymbol{B}\times\boldsymbol{C}) & \cdots & a_{nn}(\boldsymbol{B}\times\boldsymbol{C})\end{pmatrix}$$

$$=\boldsymbol{A}\times(\boldsymbol{B}\times\boldsymbol{C}).$$

定理 6 若 $\boldsymbol{A}$、$\boldsymbol{B}$ 都是对称(厄米特)矩阵,那末 $\boldsymbol{A}\times\boldsymbol{B}$ 也是对称(厄米特)的.

证明 由 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 之对称性,据定理 1 得知 $(\boldsymbol{A}\times\boldsymbol{B})'=\boldsymbol{A}'\times\boldsymbol{B}'=\boldsymbol{A}\times\boldsymbol{B}$, 即证明了 $\boldsymbol{A}\times\boldsymbol{B}$ 也是对称的. 若 $\overline{\boldsymbol{A}}'=\boldsymbol{A}$, $\overline{\boldsymbol{B}}'=\boldsymbol{B}$, 则 $\overline{(\boldsymbol{A}\times\boldsymbol{B})}'=(\overline{\boldsymbol{A}}\times\overline{\boldsymbol{B}})'=\overline{\boldsymbol{A}}'\times\overline{\boldsymbol{B}}'=\boldsymbol{A}\times\boldsymbol{B}$, 即证明了 $\boldsymbol{A}\times\boldsymbol{B}$ 为厄米特的.

定理 7 设 $\boldsymbol{A}$ 之 n 个特征值是 λ_1、λ_2、…、λ_n, $\boldsymbol{B}$ 之 m 个特征值为 μ_1、μ_2、…、μ_m, 那末 $\boldsymbol{A}\times\boldsymbol{B}$ 的 mn 个特征值是 $\lambda_i\mu_j(i=1, \cdots, n;\ j=1, \cdots, m)$.

证明 存在满秩矩阵 $\boldsymbol{P}$ 和 $\boldsymbol{Q}$, 使

$$\boldsymbol{A}_1=\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P} \quad 与 \quad \boldsymbol{B}_1=\boldsymbol{Q}^{-1}\boldsymbol{B}\boldsymbol{Q}$$

分别为 $\boldsymbol{A}$, $\boldsymbol{B}$ 之约旦标准形, 如

$$\boldsymbol{A}_1=\begin{pmatrix}\lambda_1 & & *\\ & \lambda_2 & \\ & & \ddots & \\ 0 & & & \lambda_n\end{pmatrix},\quad \boldsymbol{B}_1=\begin{pmatrix}\mu_1 & & *\\ & \mu_2 & \\ & & \ddots & \\ & & & \mu_m\end{pmatrix},$$

由定理 2, 有

$$\begin{aligned}A\times B&=(PA_1P^{-1})\times(QB_1Q^{-1})\\&=(PA_1\times QB_1)\cdot(P^{-1}\times Q^{-1})\\&=(P\times Q)\cdot(A_1\times B_1)\cdot(P^{-1}\times Q^{-1}),\end{aligned}$$

并又有

$$(P\times Q)\cdot(P^{-1}\times Q^{-1})=PP^{-1}\times QQ^{-1}=E_n\times E_m=E_{mn},$$

后者说明了$(P\times Q)^{-1}=P^{-1}\times Q^{-1}$,因之上一式为

$$A\times B=(P\times Q)\cdot(A_1\times B_1)\cdot(P\times Q)^{-1},$$

即 $A\times B\sim A_1\times B_1$, 故 $A\times B$ 与 $A_1\times B_1$ 有相同的特征值. 但由定理 4 知 $A_1\times B_1$ 为上三角矩阵, 并易看出其主对角线上元素为 $\lambda_i\mu_j(i=1, 2, \cdots, n; j=1, 2, \cdots, m)$, 故 $A\times B$ 之特征值为 $\lambda_i\mu_j$. 证完.

又, 在上证明过程中, 若取行列式, 则有

$$\begin{aligned}|A\times B|&=|P\times Q|\cdot|A_1\times B_1|\cdot|P\times Q|^{-1}=|A_1\times B_1|\\&=\prod_{j=1}^{m}(\lambda_1\mu_j)\cdot\prod_{j=1}^{m}(\lambda_2\mu_j)\cdots\prod_{j=1}^{m}(\lambda_n\mu_j)\\&=(\lambda_1\lambda_2\cdots\lambda_n)^m\cdot\left(\prod_{j=1}^{m}\mu_j\right)^n=|A|^m|B|^n,\end{aligned}$$

故又得:

定理 8 设 A 为 n 级矩阵, B 为 m 级矩阵, 则 $|A\times B|=|A|^m\cdot|B|^n$.

由定理 8 以及定理 7 之证明过程中, 又有

定理 9 $A\times B$ 为满秩的充要条件是 A 与 B 都是满秩的, 并且 $(A\times B)^{-1}=A^{-1}\times B^{-1}$.

这样, 当 A 与 B 都是直交矩阵时, 有

$$(A\times B)'=A'\times B'=A^{-1}\times B^{-1}=(A\times B)^{-1};$$

而当 A 与 B 都是酉矩阵时, 就有

$$\begin{aligned}\overline{(A\times B)}'&=(\overline{A}\times\overline{B})'=\overline{A}'\times\overline{B}'=A^{-1}\times B^{-1}\\&=(A\times B)^{-1}.\end{aligned}$$

故又有：

定理 10 设 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 都是直交矩阵(或酉矩阵)，则 $\boldsymbol{A}\times\boldsymbol{B}$ 也是直交矩阵(或酉矩阵).

定理 9 是说明：当且仅当 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 都是满秩时，$\boldsymbol{A}\times\boldsymbol{B}$ 为满秩的. 这时 $\boldsymbol{A}\times\boldsymbol{B}$ 之秩等于 $\boldsymbol{A}$ 之秩 n 与 $\boldsymbol{B}$ 之秩 m 的乘积 mn. 一般地说，有 $r_{\boldsymbol{A}\times\boldsymbol{B}}=r_{\boldsymbol{A}}\cdot r_{\boldsymbol{B}}$. 事实上，因有满秩矩阵 $\boldsymbol{H}$, $\boldsymbol{K}$, $\boldsymbol{L}$, $\boldsymbol{M}$ 使

$$\boldsymbol{A}=\boldsymbol{H}\boldsymbol{A}_1\boldsymbol{K},\quad \boldsymbol{B}=\boldsymbol{L}\boldsymbol{B}_1\boldsymbol{M},$$

这里

$$\boldsymbol{A}_1=\operatorname{diag}\{\underbrace{1,\ 1,\ \cdots,\ 1}_{r_A\text{个}},\ 0,\ \cdots,\ 0\},$$

$$\boldsymbol{B}_1=\operatorname{diag}\{\underbrace{1,\ \cdots,\ 1}_{r_B\text{个}},\ 0,\ \cdots,\ 0\}.$$

再利用定理 2，得知

$$\begin{aligned}\boldsymbol{A}\times\boldsymbol{B}&=\boldsymbol{H}\boldsymbol{A}_1\boldsymbol{K}\times\boldsymbol{L}\boldsymbol{B}_1\boldsymbol{M}\\&=(\boldsymbol{H}\times\boldsymbol{L})\cdot(\boldsymbol{A}_1\times\boldsymbol{B}_1)\cdot(\boldsymbol{K}\times\boldsymbol{M}),\end{aligned}$$

故由 $\boldsymbol{H}\times\boldsymbol{L}$ 与 $\boldsymbol{K}\times\boldsymbol{M}$ 是满秩的，可知 $r_{(\boldsymbol{A}\times\boldsymbol{B})}=r_{(\boldsymbol{A}_1\times\boldsymbol{B}_1)}$. 但 $\boldsymbol{A}_1\times\boldsymbol{B}_1$ 是 mn 级的对角矩阵，其主对角线上(由直积之定义)恰有 $r_{\boldsymbol{A}}\cdot r_{\boldsymbol{B}}$ 个元为 1，而其它元全是零，因之 $r_{(\boldsymbol{A}_1\times\boldsymbol{B}_1)}=r_{\boldsymbol{A}}r_{\boldsymbol{B}}$. 故又得：

定理 11 $r_{(\boldsymbol{A}\times\boldsymbol{B})}=r_{\boldsymbol{A}}\cdot r_{\boldsymbol{B}}$(即 $\boldsymbol{A}\times\boldsymbol{B}$ 之秩等于 $\boldsymbol{A}$ 之秩与 $\boldsymbol{B}$ 之秩的积).

前面已说过，$\boldsymbol{A}\times\boldsymbol{B}$ 与 $\boldsymbol{B}\times\boldsymbol{A}$ 一般不相等，即 $\boldsymbol{A}\times\boldsymbol{B}\neq\boldsymbol{B}\times\boldsymbol{A}$. 但 $\boldsymbol{A}\times\boldsymbol{B}$ 与 $\boldsymbol{B}\times\boldsymbol{A}$ 有怎样的关系呢？据定理 7，可知 $\boldsymbol{A}\times\boldsymbol{B}$ 与 $\boldsymbol{B}\times\boldsymbol{A}$ 有相同的特征值，因之它们有相同的特征多项式. 此外，还有别的关系吗？

由定理 2，可知 $\boldsymbol{A}\times\boldsymbol{B}=(\boldsymbol{A}\times\boldsymbol{E}_m)\cdot(\boldsymbol{E}_n\times\boldsymbol{B})$. 但

$$A\times E_m=\begin{pmatrix}\begin{matrix}a_{11}&&\\&\ddots&\\&&a_{11}\end{matrix}&\cdots&\begin{matrix}a_{1n}&&\\&\ddots&\\&&a_{1n}\end{matrix}\\ \cdots\cdots&\cdots&\cdots\cdots\\ \begin{matrix}a_{n1}&&\\&\ddots&\\&&a_{n1}\end{matrix}&\cdots&\begin{matrix}a_{nn}&&\\&\ddots&\\&&a_{nn}\end{matrix}\end{pmatrix},$$

调换它的行，且同时也调换它的同序的列，易使它变为

$$\begin{pmatrix}\begin{matrix}a_{11}&\cdots&a_{1n}\\\vdots&&\vdots\\a_{n1}&\cdots&a_{nn}\end{matrix}&&\\&\ddots&\\&&\begin{matrix}a_{11}&\cdots&a_{1n}\\\vdots&&\vdots\\a_{n1}&\cdots&a_{nn}\end{matrix}\end{pmatrix}=E_m\times A.$$

这样调换的结果无异乎是说有满秩矩阵 F 使

$$F^{-1}\cdot(A\times E_m)\cdot F=E_m\times A.$$

同理，把刚才从 $A\times E_m$ 变成 $E_m\times A$ 应调换的行及列的调动方法施行在矩阵 $E_n\times B$ 上，就能使

$$E_n\times B=\begin{pmatrix}\begin{matrix}b_{11}&\cdots&b_{1m}\\\vdots&&\vdots\\b_{m1}&\cdots&b_{mm}\end{matrix}&&\\&\ddots&\\&&\begin{matrix}b_{11}&\cdots&b_{1m}\\\vdots&&\vdots\\b_{m1}&\cdots&b_{mm}\end{matrix}\end{pmatrix}$$

变成
$$\left(\begin{array}{ccc}\begin{array}{|ccc|}b_{11} & & \\ & \ddots & \\ & & b_{11}\end{array} & \cdots & \begin{array}{|ccc|}b_{1m} & & \\ & \ddots & \\ & & b_{1m}\end{array} \\ \cdots\cdots & \cdots\cdots & \cdots\cdots \\ \begin{array}{|ccc|}b_{m1} & & \\ & \ddots & \\ & & b_{m1}\end{array} & \cdots & \begin{array}{|ccc|}b_{mm} & & \\ & \ddots & \\ & & b_{mm}\end{array}\end{array}\right)=\boldsymbol{B}\times\boldsymbol{E}_n.$$

因之又有 $\boldsymbol{F}^{-1}\cdot(\boldsymbol{E}_n\times\boldsymbol{B})\cdot\boldsymbol{F}=\boldsymbol{B}\times\boldsymbol{E}_n.$

于是，
$$\begin{aligned}\boldsymbol{A}\times\boldsymbol{B}&=(\boldsymbol{A}\times\boldsymbol{E}_m)\cdot(\boldsymbol{E}_n\times\boldsymbol{B})\\&=\boldsymbol{F}\cdot(\boldsymbol{E}_m\times\boldsymbol{A})\cdot\boldsymbol{F}^{-1}\cdot\boldsymbol{F}\cdot(\boldsymbol{B}\times\boldsymbol{E}_n)\cdot\boldsymbol{F}^{-1}\\&=\boldsymbol{F}\cdot(\boldsymbol{E}_m\times\boldsymbol{A})\cdot(\boldsymbol{B}\times\boldsymbol{E}_n)\cdot\boldsymbol{F}^{-1}\\&=\boldsymbol{F}\cdot(\boldsymbol{E}_m\boldsymbol{B}\times\boldsymbol{A}\boldsymbol{E}_n)\cdot\boldsymbol{F}^{-1}=\boldsymbol{F}\cdot(\boldsymbol{B}\times\boldsymbol{A})\cdot\boldsymbol{F}^{-1},\end{aligned}$$

说明了 $\boldsymbol{A}\times\boldsymbol{B}\sim\boldsymbol{B}\times\boldsymbol{A}$. 故又得

定理 12 $\boldsymbol{A}\times\boldsymbol{B}\sim\boldsymbol{B}\times\boldsymbol{A}$（因之它们有相同的特征多项式与最小多项式）.

练 习 四 十

1. 证明下列各式:

   i) $(\lambda\boldsymbol{A})\times\boldsymbol{B}=\boldsymbol{A}\times(\lambda\boldsymbol{B})=\lambda(\boldsymbol{A}\times\boldsymbol{B})$，$\lambda$ 为任何数;

   ii) $(\boldsymbol{A}+\boldsymbol{B})\times\boldsymbol{C}=\boldsymbol{A}\times\boldsymbol{C}+\boldsymbol{B}\times\boldsymbol{C}$;

   iii) $\boldsymbol{A}\times(\boldsymbol{B}+\boldsymbol{C})=\boldsymbol{A}\times\boldsymbol{B}+\boldsymbol{A}\times\boldsymbol{C}$.

2. 两个反对称(或反厄米特)矩阵之直积是对称(或厄米特)的，但对称(厄米特)与反对称(反厄米特)矩阵之直积是反对称(反厄米特)的.

3. 从满秩矩阵，说明 $(\boldsymbol{A}\times\boldsymbol{B})^*=\boldsymbol{A}^*\times\boldsymbol{B}^*$ 一般是不成立的.

第 2 节 复 合 矩 阵

给了 n 级矩阵 $\boldsymbol{A}$，我们知道有伴随矩阵 $\boldsymbol{A}^*$，它的元素是

$\boldsymbol{A}$ 之各元在行列式 $|\boldsymbol{A}|$ 中的代数余子式，所以，除去正负符号不计外，它的元素就是 $\boldsymbol{A}$ 的一切 $n-1$ 级子式．一般地，对于以 $\boldsymbol{A}$ 之一切 k 级子式当做元素按照某种方法排列所成的矩阵，即有：

定义 一个 n 级矩阵 $\boldsymbol{A}$ 中所有的 k 级子式总共有 $\binom{n}{k}^2=\left[\frac{n!}{k!(n-k)!}\right]^2$ 个，今把它们按照下面的方法排列成为一个 $\binom{n}{k}$ 级的矩阵 $\boldsymbol{A}^{(k)}$：

i) 凡由 $\boldsymbol{A}$ 之某 k 个行所组成的 $\binom{n}{k}$ 个 k 级子式为新矩阵中同一行的元素；

ii) 凡由 $\boldsymbol{A}$ 之某 k 个列所组成的 $\binom{n}{k}$ 个 k 级子式为新矩阵中同一列的元素；

iii) 新矩阵之第一、二、…一直到第 $\binom{n}{k}$ 行（列）都是依照辞典式的次序来排列的．

为了说明条件 iii) 中"依照辞典式的次序来排列"这句话的意义，今举一例如下．设 $\boldsymbol{A}$ 为 4 级矩阵，则由 $\boldsymbol{A}$ 之第一、三两行所组成的一切 2 级子式必须排在由 $\boldsymbol{A}$ 之第二、三两行，或第一、四两行，或第二、四两行，或第三、四两行各种情形所组成的一切 2 级子式的前面．同样，对于列也是如此．因之，如果

$$\boldsymbol{A}=\begin{pmatrix}a_{11}&a_{12}&a_{13}&a_{14}\\a_{21}&a_{22}&a_{23}&a_{24}\\a_{31}&a_{32}&a_{33}&a_{34}\\a_{41}&a_{42}&a_{43}&a_{44}\end{pmatrix},$$

则

$$A^{(2)}=\begin{pmatrix} \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} & \begin{vmatrix} a_{11} & a_{13} \\ a_{21} & a_{23} \end{vmatrix} & \cdots & \begin{vmatrix} a_{12} & a_{14} \\ a_{22} & a_{24} \end{vmatrix} & \begin{vmatrix} a_{13} & a_{14} \\ a_{23} & a_{24} \end{vmatrix} \\ \begin{vmatrix} a_{11} & a_{12} \\ a_{31} & a_{32} \end{vmatrix} & \begin{vmatrix} a_{11} & a_{13} \\ a_{31} & a_{33} \end{vmatrix} & \cdots & \begin{vmatrix} a_{12} & a_{14} \\ a_{32} & a_{34} \end{vmatrix} & \begin{vmatrix} a_{13} & a_{14} \\ a_{33} & a_{34} \end{vmatrix} \\ \vdots & \vdots & & \vdots & \vdots \\ \begin{vmatrix} a_{31} & a_{32} \\ a_{41} & a_{42} \end{vmatrix} & \begin{vmatrix} a_{31} & a_{33} \\ a_{41} & a_{43} \end{vmatrix} & \cdots & \begin{vmatrix} a_{32} & a_{34} \\ a_{42} & a_{44} \end{vmatrix} & \begin{vmatrix} a_{33} & a_{34} \\ a_{43} & a_{44} \end{vmatrix} \end{pmatrix}.$$

象在定义中这样所形成的 $\binom{n}{k}$ 级矩阵 $A^{(k)}$，叫做 A 之复合矩阵，更详细或确切些叫做 A 的第 k 次复合矩阵.

从复合矩阵之定义，易知：$A^{(k)}$ 的主对角线上的元素都是 A 之 k 级主子式，不在 $A^{(k)}$ 之主对角线上的元素都不是 A 之 k 级主子式.

又，从 $A^{(k)}$ 之定义，也易知 $A^{(1)}=A$，$A^{(n)}=|A|$.

关于单位矩阵、对角矩阵、纯量矩阵的复合矩阵，有：

定理 1 i) $E_n^{(k)}=E_{\binom{n}{k}}$——即 n 级单位矩阵的第 k 次复合矩阵为 $\binom{n}{k}$ 级的单位矩阵；

ii) 若 $D=\mathrm{diag}\{\lambda_1, \lambda_2, \cdots, \lambda_n\}$，则 $D^{(k)}=\mathrm{diag}\{p_1, p_2, \cdots, p_{\binom{n}{k}}\}$，其中 p_i 为 k 个 λ_j 的乘积；

iii) 若 $D=cE_n$，则 $D^{(k)}=c^kE_{\binom{n}{k}}$.

这都可由定义直接推得，读者自行写出证明.

现在来研究 $A^{(k)}$ 的性质、首先问 $(A^{(k)})'=?$

为明显起见，令

$$
A=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix},
$$

则
$$
A'=\begin{pmatrix} a_{11} & a_{21} & \cdots & a_{n1} \\ a_{12} & a_{22} & \cdots & a_{n2} \\ \vdots & \vdots & & \vdots \\ a_{1n} & a_{2n} & \cdots & a_{nn} \end{pmatrix}.
$$

若 $A^{(k)}$ 之第 i 行第 j 列处的元素为

$$
(A^{(k)})_{ij}=d(A^{i_1i_2\cdots i_k}_{j_1j_2\cdots j_k}),
$$

则
$$
(A^{(k)})'_{ji}=(A^{(k)})_{ij}=d(A^{i_1i_2\cdots i_k}_{j_1j_2\cdots j_k})
$$

$$
=\begin{vmatrix} a_{i_1j_1} & a_{i_1j_2} & \cdots & a_{i_1j_k} \\ a_{i_2j_1} & a_{i_2j_2} & \cdots & a_{i_2j_k} \\ \cdots & \cdots & \cdots & \cdots \\ a_{i_kj_1} & a_{i_kj_2} & \cdots & a_{i_kj_k} \end{vmatrix}
$$

$$
=\begin{vmatrix} a_{i_1j_1} & a_{i_2j_1} & \cdots & a_{i_kj_1} \\ a_{i_1j_2} & a_{i_2j_2} & \cdots & a_{i_kj_2} \\ \cdots & \cdots & \cdots & \cdots \\ a_{i_1j_k} & a_{i_2j_k} & \cdots & a_{i_kj_k} \end{vmatrix}
$$

$$
=d(A'^{j_1j_2\cdots j_k}_{i_1i_2\cdots i_k})=(A'^{(k)})_{ji},
$$

即说明了两矩阵 $(A^{(k)})'$ 与 $(A')^{(k)}$ 之对应位置之元都相等，故得:

定理 2 $(A^{(k)})'=(A')^{(k)}$.

这样，当 A 是厄米特矩阵时(即 $\bar{A}'=A$)，因 $\overline{(A^{(k)})}=\bar{A}^{(k)}$，故 $\overline{(A^{(k)})}'=(\bar{A}^{(k)})'=(\bar{A}')^{(k)}=A^{(k)}$，说明了 $A^{(k)}$ 是厄米特的. 同样，若 A 为对称的，则 $A^{(k)}$ 也对称. 但当 A 为反厄米特矩阵时($\bar{A}'=-A$)，则有

$$\overline{(A^{(k)})'} = (\overline{A}^{(k)})' = (\overline{A}')^{(k)} = (-A)^{(k)} = (-1)^k A^{(k)}$$

$$= \begin{cases} A^{(k)}, & \text{当 } k \text{ 是偶数时;} \\ -A^{(k)}, & \text{当 } k \text{ 是奇数时.} \end{cases}$$

对于反对称矩阵,也有类似的结果. 这就证明了:

定理 3 若 A 是厄米特(对称)矩阵,则 $A^{(k)}$ 也是厄米特(对称)矩阵. 若 A 是反厄米特(反对称)矩阵,则 $A^{(k)}$ 是厄米特或反厄米特(对称或反对称),完全由 k 为偶数或奇数来决定.

至于 A 为直交矩阵或酉矩阵时, $A^{(k)}$ 又怎样呢? 再问 $A^{(k)}$ 之秩与 A 之秩间有什么关系呢? 这些问题,都有赖于两个矩阵之积的复合矩阵之构造.

设 $P=AB$. 若 $P^{(k)}$ 之第 i 行第 j 列处的元素是

$$(P^{(k)})_{ij} = d(P^{i_1 i_2 \cdots i_k}_{j_1 j_2 \cdots j_k}),$$

则据复合矩阵之定义,即知 $A^{(k)}$ 之第 i 行各元素都是 $d(A^{i_1 i_2 \cdots i_k}_{r_1 r_2 \cdots r_k})$ 的形状, $B^{(k)}$ 之第 j 列的诸元素又都是 $d(B^{r_1 r_2 \cdots r_k}_{j_1 j_2 \cdots j_k})$ 的形状; 且当前者为 $A^{(k)}$ 之第 i 行第 t 列处之元时,后者就一定是 $B^{(k)}$ 之第 t 行第 j 列处的元. 但从 $P=AB$ 又知

$$d(P^{i_1 i_2 \cdots i_k}_{j_1 j_2 \cdots j_k}) = \sum_{(r_1, r_2, \cdots, r_k)} d(A^{i_1 i_2 \cdots i_k}_{r_1 r_2 \cdots r_k}) \cdot d(B^{r_1 r_2 \cdots r_k}_{j_1 j_2 \cdots j_k}),$$

而右边之 $\sum$ 显然是 $A^{(k)}B^{(k)}$ 之第 i 行第 j 列处的元 $(A^{(k)}B^{(k)})_{ij}$, 即 $(P^{(k)})_{ij} = ((AB)^{(k)})_{ij} = (A^{(k)}B^{(k)})_{ij}$, 故

$$P^{(k)} = (AB)^{(k)} = A^{(k)}B^{(k)}.$$

再由归纳法,可知上式对任意多个矩阵也成立. 于是证得了:

定理 4 设 $A_1, A_2, \cdots, A_t$ 为同级的矩阵,则有

$$(A_1 A_2 \cdots A_t)^{(k)} = A_1^{(k)} A_2^{(k)} \cdots A_t^{(k)}.$$

这样,当 n 级矩阵 A 是满秩时,从 $AA^{-1} = E_n$, 即得

$$E_{\binom{n}{k}} = E_n^{(k)} = (AA^{-1})^{(k)} = A^{(k)} \cdot (A^{-1})^{(k)},$$

故又得：

推论 1 若 $\boldsymbol{A}$ 为满秩的，则 $\boldsymbol{A}^{(k)}$ 也是满秩的，并且 $(\boldsymbol{A}^{(k)})^{-1}=(\boldsymbol{A}^{-1})^{(k)}$.

一般，若 $\boldsymbol{A}$ 之秩为 r，则有二个满秩阵 $\boldsymbol{P}$ 与 $\boldsymbol{Q}$ 使

$$\boldsymbol{PAQ}=\boldsymbol{D}=\mathrm{diag}\,\{\underbrace{1,\ \cdots,\ 1}_{r\text{个}},\ 0,\ \cdots,\ 0\},$$

因之 $\boldsymbol{D}^{(k)}=\boldsymbol{P}^{(k)}\boldsymbol{A}^{(k)}\boldsymbol{Q}^{(k)}$；然由上推论 1，$\boldsymbol{P}^{(k)}$ 与 $\boldsymbol{Q}^{(k)}$ 是满秩的，故 $r_{\boldsymbol{D}^{(k)}}=r_{\boldsymbol{A}^{(k)}}$. 据定理 1，$\boldsymbol{D}^{(k)}$ 是对角阵，其主对角线上之元是 $\boldsymbol{D}$ 中 k 个对角线上元素的积，故当 $k>r$ 时有 $\boldsymbol{D}^{(k)}=\boldsymbol{O}$，因而 $\boldsymbol{A}^{(k)}=\boldsymbol{O}$，但在 $k\leqslant r$ 时，$\boldsymbol{D}^{(k)}$ 之对角线上不为零的数都是从 r 个 1 中取 k 个，因而其个数显为

$$\binom{r}{k}=\frac{r!}{k!\,(r-k)!},$$

即
$$\boldsymbol{D}^{(k)}\text{之秩}=\binom{r}{k},$$

随而
$$\boldsymbol{A}^{(k)}\text{之秩}=\binom{r}{k}.$$

故又得：

推论 2 *若 $\boldsymbol{A}$ 之秩 $=r$，则当 $k\leqslant r$ 时，$\boldsymbol{A}^{(k)}$ 之秩为 $\binom{r}{k}$，而在 $k>r$ 时，$\boldsymbol{A}^{(k)}=\boldsymbol{O}$.*

注意：1. 由复合矩阵之定义，易知推论 2 中 $\boldsymbol{A}^{(k)}=\boldsymbol{O}$（当 $k>r$ 时），因这时 $\boldsymbol{A}$ 中任何 k 级子式（$k>r$）都是零.

2. 特别地，当 $\boldsymbol{A}$ 是满秩时（$r=n$），由推论 2 知 $r_{\boldsymbol{A}^{(k)}}=\binom{n}{k}$，即 $\boldsymbol{A}^{(k)}$ 是满秩的，这说明了推论 1 之前半是推论 2 的特例. 可是，应注意的是，在证明推论 2 时，还要以推论 1 为依据.

推论 1 只是说当 $\boldsymbol{A}$ 为满秩时，$\boldsymbol{A}^{(k)}$ 也是满秩的. 至于 $\boldsymbol{A}^{(k)}$ 是满秩时，$\boldsymbol{A}$ 之秩怎样呢？实际上，据推论 2，当 $\boldsymbol{A}$ 之秩 $=r$ 时，$\boldsymbol{A}^{(k)}$ 之秩 $=\binom{r}{k}$，$k\leqslant r\leqslant n$，故从 $\boldsymbol{A}^{(k)}$ 之满秩性，知其秩 $=\binom{n}{k}$，因而这时有 $\binom{n}{k}=\binom{r}{k}$，即

$$\frac{n!}{k!(n-k)!}=\frac{r!}{k!(r-k)!},\quad \frac{n!}{r!}=\frac{(n-k)!}{(r-k)!},$$

故 $\quad n(n-1)\cdots(r+1)=(n-k)(n-1-k)\cdots(r+1-k)$.

于是，若 $n>r$，则因上式之左右两边都有 $n-r$ 个因子，并且

$$n>n-k,\ n-1>n-1-k,\ \cdots,\ r+1>r+1-k,$$

故又有

$$n(n-1)\cdots(r+1)>(n-k)(n-1-k)\cdots(r+1-k),$$

显与上式矛盾. 因而只能是 $n=r$，即 $\boldsymbol{A}$ 是满秩的. 于是，再与推论 1 合并，又得到了：

定理 5 $\boldsymbol{A}^{(k)}$ 为满秩的充要条件是 $\boldsymbol{A}$ 为满秩的.

由推论 2 可知，当 $\boldsymbol{A}$ 之秩 $=r$ 时，$\boldsymbol{A}^{(r)}$ 之秩 $=\binom{r}{r}=1$，因之 $\boldsymbol{A}^{(r)}$ 中任意二级子式等于零，即

$$\begin{vmatrix} d\left(\boldsymbol{A}_{j_1j_2\cdots j_r}^{i_1i_2\cdots i_r}\right) & d\left(\boldsymbol{A}_{i_1i_2\cdots i_r}^{i_1i_2\cdots i_r}\right) \\ d\left(\boldsymbol{A}_{j_1j_2\cdots j_r}^{j_1j_2\cdots j_r}\right) & d\left(\boldsymbol{A}_{i_1i_2\cdots i_r}^{j_1j_2\cdots j_r}\right) \end{vmatrix}=0,$$

或

$$d\left(\boldsymbol{A}_{j_1j_2\cdots j_r}^{i_1i_2\cdots i_r}\right)\cdot d\left(\boldsymbol{A}_{i_1i_2\cdots i_r}^{j_1j_2\cdots j_r}\right)=d\left(\boldsymbol{A}_{i_1i_2\cdots i_r}^{i_1i_2\cdots i_r}\right)\cdot d\left(\boldsymbol{A}_{j_1j_2\cdots j_r}^{j_1j_2\cdots j_r}\right). \tag{1}$$

于是，若 $\boldsymbol{A}$ 为实对称时，易知

$$d\left(\boldsymbol{A}_{j_1j_2\cdots j_r}^{i_1i_2\cdots i_r}\right)=d\left(\boldsymbol{A}_{i_1i_2\cdots i_r}^{j_1j_2\cdots j_r}\right)$$

（因它们互为转置的行列式），故由 (1) 即得

$$d\left(\boldsymbol{A}_{i_1i_2\cdots i_r}^{i_1i_2\cdots i_r}\right)\cdot d\left(\boldsymbol{A}_{j_1j_2\cdots j_r}^{j_1j_2\cdots j_r}\right)=\left[d\left(\boldsymbol{A}_{j_1j_2\cdots j_r}^{i_1i_2\cdots i_r}\right)\right]^2.$$

由 $r_A=r$，知必有某 $d\left(\boldsymbol{A}_{j_1j_2\cdots j_r}^{i_1i_2\cdots i_r}\right)\neq 0$，因而据上式，有

$$d\left(\boldsymbol{A}_{i_1i_2\cdots i_r}^{i_1i_2\cdots i_r}\right)\cdot d\left(\boldsymbol{A}_{j_1j_2\cdots j_r}^{j_1j_2\cdots j_r}\right)>0.$$

如 $\boldsymbol{A}$ 为厄米特的，则

$$d(\boldsymbol{A}^{i_1 i_2\cdots i_r}_{j_1 j_2\cdots j_r})=\overline{d(\boldsymbol{A}^{j_1 j_2\cdots j_r}_{i_1 i_2\cdots i_r})},$$

因而(1)式则变为

$$d(\boldsymbol{A}^{i_1 i_2\cdots i_r}_{i_1 i_2\cdots i_r})\cdot d(\boldsymbol{A}^{j_1 j_2\cdots j_r}_{j_1 j_2\cdots j_r})=d(\boldsymbol{A}^{i_1 i_2\cdots i_r}_{j_1 j_2\cdots j_r})\cdot\overline{d(\boldsymbol{A}^{i_1 i_2\cdots i_r}_{j_1 j_2\cdots j_r})}>0,$$

当 $d(\boldsymbol{A}^{i_1 i_2\cdots i_r}_{j_1 j_2\cdots j_r})\neq 0$ 时.

这就证明了：

定理 6 厄米特(或实对称)矩阵之秩如为 r，则它必至少有一个 r 级的主子式不等于零．且任两个非零的 r 级主子式必有相同符号(即：或皆为正数，或皆为负数)．

如果 $\boldsymbol{A}$ 是反厄米特的，容易验证：

$$d(\boldsymbol{A}^{i_1 i_2\cdots i_r}_{j_1 j_2\cdots j_r})=(-1)^r\cdot\overline{d(\boldsymbol{A}^{j_1 j_2\cdots j_r}_{i_1 i_2\cdots i_r})},$$

$$d(\boldsymbol{A}^{i_1 i_2\cdots i_r}_{i_1 i_2\cdots i_r})=(-1)^r\cdot\overline{d(\boldsymbol{A}^{i_1 i_2\cdots i_r}_{i_1 i_2\cdots i_r})}.$$

这说明了反厄米特矩阵之 r 级主子式为实数或纯虚数，全由 r 为偶数或奇数来决定．若又 $r_A=r$ 时，并设 $d(\boldsymbol{A}^{i_1 i_2\cdots i_r}_{j_1 j_2\cdots j_r})\neq 0$，则从(1)式又得

$$d(\boldsymbol{A}^{i_1 i_2\cdots i_r}_{i_1 i_2\cdots i_r})\cdot d(\boldsymbol{A}^{j_1 j_2\cdots j_r}_{j_1 j_2\cdots j_r})=(-1)^r\cdot d(\boldsymbol{A}^{i_1 i_2\cdots i_r}_{j_1 j_2\cdots j_r})\cdot\overline{d(\boldsymbol{A}^{i_1 i_2\cdots i_r}_{j_1 j_2\cdots j_r})},$$

故 $d(\boldsymbol{A}^{i_1 i_2\cdots i_r}_{i_1 i_2\cdots i_r})$ 与 $d(\boldsymbol{A}^{j_1 j_2\cdots j_r}_{j_1 j_2\cdots j_r})$ 在 r 为偶数时是具同符号的异于零之实数，而在 r 为奇数时是虚部具同符号的异于零之纯虚数．故又得

定理 7 反厄米特矩阵之秩若为 r，则它至少有一个 r 级主子式不等于零；且任两个非零的 r 级主子式在 r 为偶数时是具同符号的实数，而在 r 为奇数时是虚部具同符号的纯虚数．

若 $\boldsymbol{A}$ 为实反对称的，则知：

$$d(\boldsymbol{A}^{i_1 i_2\cdots i_r}_{j_1 j_2\cdots j_r})=(-1)^r\cdot d(\boldsymbol{A}^{j_1 j_2\cdots j_r}_{i_1 i_2\cdots i_r}),$$

因而

$$d(\boldsymbol{A}^{i_1 i_2\cdots i_r}_{i_1 i_2\cdots i_r})=(-1)^r\cdot d(\boldsymbol{A}^{i_1 i_2\cdots i_r}_{i_1 i_2\cdots i_r}).$$

这说明了 $\boldsymbol{A}$ 中奇级主子式恒为零．故若 $r_A=r$ 时，则当

$d\left(\boldsymbol{A}_{j_1 j_2 \cdots j_r}^{i_1 i_2 \cdots i_r}\right) \neq 0$, 就从(1)式应有

$$d\left(\boldsymbol{A}_{i_1 i_2 \cdots i_r}^{i_1 i_2 \cdots i_r}\right) \cdot d\left(\boldsymbol{A}_{j_1 j_2 \cdots j_r}^{j_1 j_2 \cdots j_r}\right) = (-1)^r \cdot \left[d\left(\boldsymbol{A}_{j_1 j_2 \cdots j_r}^{i_1 i_2 \cdots i_r}\right)\right]^2 \neq 0,$$

即主子式 $d\left(\boldsymbol{A}_{i_1 i_2 \cdots i_r}^{i_1 i_2 \cdots i_r}\right)$ 与 $d\left(\boldsymbol{A}_{j_1 j_2 \cdots j_r}^{j_1 j_2 \cdots j_r}\right)$ 都不等于零, 故有 r 为偶数, 因之 $d\left(\boldsymbol{A}_{i_1 i_2 \cdots i_r}^{i_1 i_2 \cdots i_r}\right)$ 与 $d\left(\boldsymbol{A}_{j_1 j_2 \cdots j_r}^{j_1 j_2 \cdots j_r}\right)$ 同号. 即又得:

定理 8 实反对称矩阵之秩 r 必为偶数, 且必有一 r 级主子式不等于零. 又, 凡非零的 r 级主子式皆同号.

再, 若 $\boldsymbol{A}$ 为酉矩阵, 则 $\boldsymbol{A}\overline{\boldsymbol{A}}' = \boldsymbol{E}_n$; 利用定理 4 与定理 2, 就有

$$\boldsymbol{E}_n^{(k)} = \boldsymbol{A}^{(k)} \cdot (\overline{\boldsymbol{A}}')^{(k)} = \boldsymbol{A}^{(k)} \cdot (\overline{\boldsymbol{A}^{(k)}})' = \boldsymbol{A}^{(k)} \cdot \overline{(\boldsymbol{A}^{(k)})}',$$

即

$$\boldsymbol{A}^{(k)} \cdot \overline{(\boldsymbol{A}^{(k)})}' = \boldsymbol{E}_{\binom{n}{k}}.$$

对直交矩阵 $\boldsymbol{A}$, 也有

$$\boldsymbol{A}^{(k)} \cdot \overline{(\boldsymbol{A}^{(k)})}' = \boldsymbol{E}_{\binom{n}{k}}.$$

于是又证得了:

定理 9 酉矩阵(直交矩阵)之复合矩阵还是酉矩阵(直交矩阵).

再研究三角矩阵. 只就上三角矩阵 $\boldsymbol{A} = (a_{ij})$ 来考虑(对于下三角矩阵, 也可类似地讨论), 于是 $a_{ij} = 0\ (i > j)$. 故若令 $\boldsymbol{A}^{(k)} = (c_{ij})$, 其中

$$c_{ij} = d\left(\boldsymbol{A}_{j_1 j_2 \cdots j_k}^{i_1 i_2 \cdots i_k}\right), \quad \begin{pmatrix} i_1 < i_2 < \cdots < i_k, \\ j_1 < j_2 < \cdots < j_k. \end{pmatrix}$$

那末, 当 $i > j$ 时由定义已知 c_{ij} 不是 $\boldsymbol{A}$ 的 k 级主子式, 因而 k 个等式 $i_1 = j_1$, $i_2 = j_2$, $\cdots$, $i_k = j_k$ 不可同时成立. 今假定

$$i_1 = j_1, \ \cdots, \ i_{t-1} = j_{t-1}, \ i_t \neq j_t, \quad (1 \leqslant t \leqslant k),$$

则因 $i > j$, 故由复合矩阵之定义, 必有 $i_t > j_t$. 再利用 $a_{\lambda\mu} = 0$ $(\lambda > \mu)$, 可得

$$c_{ij}=d\left(A^{i_1 i_2\cdots i_k}_{j_1 j_2\cdots j_k}\right)=\begin{vmatrix} a_{i_1 j_1} & \cdots & a_{i_1 j_t} & a_{i_1 j_{t+1}} & \cdots & a_{i_1 j_k} \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{i_t j_1} & \cdots & a_{i_t j_t} & a_{i_t j_{t+1}} & \cdots & a_{i_t j_k} \\ 0 & \cdots & 0 & a_{i_{t+1} j_{t+1}} & \cdots & a_{i_{t+1} j_k} \\ \vdots & & \vdots & \vdots & & \vdots \\ 0 & \cdots & 0 & a_{i_k j_{t+1}} & \cdots & a_{i_k j_k} \end{vmatrix}$$

$$=\begin{vmatrix} a_{i_1 j_1} & \cdots & a_{i_1 j_t} \\ \vdots & & \vdots \\ a_{i_t j_1} & \cdots & a_{i_t j_t} \end{vmatrix}\cdot\begin{vmatrix} a_{i_{t+1} j_{t+1}} & \cdots & a_{i_{t+1} j_k} \\ \vdots & & \vdots \\ a_{i_k j_{t+1}} & \cdots & a_{i_k j_k} \end{vmatrix}=0$$

(因 $a_{i_t j_1}=\cdots=a_{i_t j_t}=0$). 此即表示这时 $\boldsymbol{A}^{(k)}$ 也是上三角矩阵. 又因 $\boldsymbol{A}^{(k)}$ 之主对角线上每元都是 $\boldsymbol{A}$ 的 k 级主子式, 故它是 $\boldsymbol{A}$ 之主对角线上 k 个元素的乘积. 于是又证得了:

定理 10 三角矩阵 $\boldsymbol{A}$ 之复合矩阵 $\boldsymbol{A}^{(k)}$ 也是同型的三角矩阵, 其主对角线每个元素等于 $\boldsymbol{A}$ 之主对角线上 k 个元素的乘积.

有了定理 10, 容易解决 $\boldsymbol{A}^{(k)}$ 的特征值与 $\boldsymbol{A}$ 的特征值之间的关系. 事实上, 有满秩矩阵 $\boldsymbol{P}$, 使 $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}=\boldsymbol{D}$, 但

$$\boldsymbol{D}=\begin{pmatrix} \lambda_1 & & * \\ & \lambda_2 & \\ & & \ddots \\ 0 & & & \lambda_n \end{pmatrix}$$

为上三角矩阵(例如 $\boldsymbol{D}$ 为 $\boldsymbol{A}$ 之约旦标准形), 于是 $\lambda_1, \lambda_2, \cdots, \lambda_n$ 是 $\boldsymbol{A}$ 的 n 个特征值. 于是,

$$\begin{aligned}\boldsymbol{D}^{(k)}&=(\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P})^{(k)}=(\boldsymbol{P}^{-1})^{(k)}\boldsymbol{A}^{(k)}\boldsymbol{P}^{(k)}\\&=(\boldsymbol{P}^{(k)})^{-1}\cdot\boldsymbol{A}^{(k)}\cdot\boldsymbol{P}^{(k)}\end{aligned}$$

说明了 $\boldsymbol{A}^{(k)}\sim\boldsymbol{D}^{(k)}$, 故由 $\boldsymbol{D}^{(k)}$ 之上三角性及其主对角线上各元等于从 $\lambda_1, \lambda_2, \cdots, \lambda_n$ 中任取 k 个之积(定理 10), 即知 $\boldsymbol{A}^{(k)}$ 之特征值等于从 $\lambda_1, \lambda_2, \cdots, \lambda_n$ 中取任 k 个之积, 也就是证

得了：

定理 11 $A^{(k)}$ 之 $\binom{n}{k}$ 个特征值中每一个都等于从 A 之 n 个特征值中取 k 个的乘积.

再据这定理 11，易求 $A^{(k)}$ 之行列式. 因为矩阵之行列式等于它的所有特征值的乘积，故从定理 11 即有

$$|A^{(k)}|=\prod_{i}(\lambda_{i_1}\lambda_{i_2}\cdots\lambda_{i_k}),$$

这里 $\lambda_{i_1}, \lambda_{i_2}, \cdots, \lambda_{i_k}$ 是从 $\lambda_1, \lambda_2, \cdots, \lambda_n$ 中任取 k 个的一组，而 $\prod\limits_{i}$ 表示这样一些组的乘积；因之在乘积 $\prod\limits_{i}$ 中每个 λ 出现的回数相同，都是 $\binom{n-1}{k-1}$ 次，故

$$|A^{(k)}|=(\lambda_1\lambda_2\cdots\lambda_n)^{\binom{n-1}{k-1}}=|A|^{\binom{n-1}{k-1}}.$$

即又证得：

定理 12 设 A 为 n 级矩阵，则 $|A^{(k)}|=|A|^{\binom{n-1}{k-1}}$.

练习四十一

1. 证明：

i) $[(A_1A_2\cdots A_t)^{-1}]^{(k)}=(A_t^{(k)})^{-1}\cdots(A_2^{(k)})^{-1}(A_1^{(k)})^{-1}$;

ii) $[(A_1A_2\cdots A_t)']^{(k)}=(A_t^{(k)})'\cdots(A_2^{(k)})'(A_1^{(k)})'$.

2. 若 A 之最小多项式无重根，证明 $A^{(k)}$ 之最小多项式也无重根.

3. 设 $A=\begin{pmatrix}1&1&0\\0&1&0\\0&0&0\end{pmatrix}$，求 $A^{(2)}$. 再求 A 与 $A^{(2)}$ 之最小多项式，并由此说明上题的逆定理是否成立？

4. 设 A 是正定厄米特矩阵，试证 $A^{(k)}$ 也是正定的. 若 A 是负定的，则 $A^{(k)}$ 或为正定或为负定，完全由 k 为偶数或奇数来决定.

第 3 节　特征值的估计

在微分方程的稳定性理论中，决定矩阵之特征值的实部

(或虚部)全在复平面的左半部或右半部是一个很重要的问题. 这里, 对矩阵之特征值的估计作一个简单的介绍. 为叙述简便, 今后恒用符号 $R(\lambda)$ 与 $I(\lambda)$ 分别表示复数 λ 的实部与虚部, 即若 $\lambda=a+bi$(a, b 是实数), 则

$$R(\lambda)=a, \quad I(\lambda)=b.$$

关于实矩阵之特征值, 有下面的

定理 1 设 $\boldsymbol{A}=(a_{rs})$ 是 n 级实矩阵, 并令

$$M=\max_{1\leqslant r,s\leqslant n}\frac{1}{2}|a_{rs}-a_{sr}|.$$

若 λ 是 $\boldsymbol{A}$ 之任一个特征值, 则

$$|I(\lambda)|\leqslant M\cdot\sqrt{\frac{n(n-1)}{2}}.$$

证明 令 $\boldsymbol{X}=\begin{pmatrix}x_1\\ \vdots\\ x_n\end{pmatrix}$ 为 $\boldsymbol{A}$ 对应于 λ 的一个特征向量, 即 $\boldsymbol{AX}=\lambda\boldsymbol{X}$. 不失一般性, 可限制 $\overline{\boldsymbol{X}}'\boldsymbol{X}=\sum\bar{x}_rx_r=1$.

$$\overline{\boldsymbol{X}}'\boldsymbol{AX}=\lambda\overline{\boldsymbol{X}}'\boldsymbol{X}=\lambda\Rightarrow\bar{\lambda}=\overline{(\overline{\boldsymbol{X}}'\boldsymbol{AX})}'=\overline{\boldsymbol{X}}'\overline{\boldsymbol{A}}'\boldsymbol{X}=\overline{\boldsymbol{X}}'\boldsymbol{A}'\boldsymbol{X},$$

故
$$\lambda-\bar{\lambda}=2i\cdot I(\lambda)=\overline{\boldsymbol{X}}'(\boldsymbol{A}-\boldsymbol{A}')\boldsymbol{X}$$

为纯虚数. 但

$$\begin{aligned}2i\cdot I(\lambda)&=\overline{\boldsymbol{X}}'(\boldsymbol{A}-\boldsymbol{A}')\boldsymbol{X}=\sum_{r,s=1}^{n}(a_{rs}-a_{sr})\bar{x}_rx_s\\&=\frac{1}{2}\left\{\sum_{r,s=1}^{n}(a_{rs}-a_{sr})\bar{x}_rx_s+\sum_{r,s=1}^{n}(a_{sr}-a_{rs})x_r\bar{x}_s\right\}\\&=\sum_{r,s=1}^{n}(a_{rs}-a_{sr})\cdot\frac{1}{2}(\bar{x}_rx_s-x_r\bar{x}_s)\end{aligned}$$

中各项 $(a_{rs}-a_{sr})\cdot\frac{1}{2}(\bar{x}_rx_s-x_r\bar{x}_s)$ 也是纯虚数, 故取绝对值即得

$$2|I(\lambda)|\leqslant\sum_{r,s=1}^{n}\frac{1}{2}\cdot|a_{rs}-a_{sr}|\cdot|\bar{x}_r x_s-x_r\bar{x}_s|$$

$$\leqslant M\cdot\sum_{\substack{r,s=1\\r\neq s}}^{n}|\bar{x}_r x_s-x_r\bar{x}_s|. \tag{2}$$

由于任 m 个实数 $y_1, y_2, \cdots, y_m$, 恒有

$$m(y_1^2+\cdots+y_m^2)-(y_1+\cdots+y_m)^2=\sum_{1\leqslant j<k\leqslant m}(y_j-y_k)^2\geqslant 0,$$

即

$$(y_1+\cdots+y_m)^2\leqslant m(y_1^2+\cdots+y_m^2), \tag{3}$$

故利用(3), 从(2)式就得知

$$4\cdot I(\lambda)^2\leqslant M^2\cdot\Big(\sum_{\substack{r,s=1\\r\neq s}}^{n}|\bar{x}_r x_s-x_r\bar{x}_s|\Big)^2$$

$$\leqslant M^2\cdot n(n-1)\cdot\sum_{\substack{r,s=1\\r\neq s}}^{n}|\bar{x}_r x_s-x_r\bar{x}_s|^2. \tag{4}$$

另方面, 因

$$\begin{aligned}|\bar{x}_r x_s-x_r\bar{x}_s|^2&=(\bar{x}_r x_s-x_r\bar{x}_s)\cdot\overline{(\bar{x}_r x_s-x_r\bar{x}_s)}\\&=(\bar{x}_r x_s-x_r\bar{x}_s)(x_r\bar{x}_s-\bar{x}_r x_s)\\&=-(\bar{x}_r^2 x_s^2-2\bar{x}_r x_r\bar{x}_s x_s+x_r^2\bar{x}_s^2)\\&=2|x_r|^2\cdot|x_s|^2-x_r^2\bar{x}_s^2-\bar{x}_r^2 x_s^2,\end{aligned}$$

故

$$\begin{aligned}\sum_{\substack{r,s=1\\r\neq s}}^{n}|\bar{x}_r x_s-x_r\bar{x}_s|^2&=2\cdot\sum_{s,r=1}^{n}|x_r|^2\cdot|x_s|^2-2\cdot\sum_{r=1}^{n}\bar{x}_r^2\cdot\sum_{s=1}^{n}x_s^2\\&=2\cdot\sum_{r=1}^{n}|x_r|^2\cdot\sum_{s=1}^{n}|x_s|^2-2\sum_{r=1}^{n}\bar{x}_r^2\cdot\sum_{s=1}^{n}x_s^2\\&=2-2|t|^2\leqslant 2,\end{aligned}$$

这里 $t=\sum_{r=1}^{n}x_r^2$ (因为 $\sum|x_r|^2=\sum\bar{x}_r x_r=1$). 因之, 从(4)式就得到了

$$4\cdot I(\lambda)^2\leqslant M^2\cdot n(n-1)\cdot 2,\quad I(\lambda)^2\leqslant M^2\cdot\frac{n(n-1)}{2},$$

定理 1 证完.

注意：当 $\boldsymbol{A}$ 为实对称矩阵时，则 $M=0$，故 $I(\lambda)=0$，即 λ 为实数. 这从另一个角度又证明了实对称矩阵的特征值都是实数.

至于复矩阵的特征值，则有下面的

定理 2 设 $\boldsymbol{A}=(a_{rs})$ 是 n 级复矩阵，并令 $\rho=\max\limits_{1\leqslant r,s\leqslant n}|a_{rs}|$，$\sigma=\max\limits_{1\leqslant r,s\leqslant n}\frac{1}{2}|a_{rs}+\bar{a}_{sr}|$，$\tau=\max\limits_{1\leqslant r,s\leqslant n}\frac{1}{2}|a_{rs}-\bar{a}_{sr}|$. 则 $\boldsymbol{A}$ 之每个特征值 λ 必满足下列关系式：

$$|\lambda|\leqslant n\rho,\quad |R(\lambda)|\leqslant n\sigma,\quad |I(\lambda)|\leqslant n\tau.$$

证明 不论 $\boldsymbol{C}=(c_{rs})$ 为任何复矩阵，若令 $|c_{rs}|\leqslant\varkappa(r,s=1,2,\cdots,n)$，则当 $\boldsymbol{X}$ 是复单位向量$\left(\text{即}\sum\limits_{r=1}^{n}\bar{x}_r x_r=1\right)$时，就有

$$|\overline{\boldsymbol{X}}'\boldsymbol{C}\boldsymbol{X}|=\left|\sum_{r,s=1}^{n}c_{rs}\bar{x}_r x_s\right|\leqslant\varkappa\sum_{r,s=1}^{n}|x_r|\cdot|x_s|$$

$$=\varkappa\cdot\left(\sum_{r=1}^{n}|x_r|\right)^2.$$

再利用(3)式就得到了

$$|\overline{\boldsymbol{X}}'\boldsymbol{C}\boldsymbol{X}|\leqslant\varkappa n\cdot\sum_{r=1}^{n}|x_r|^2=\varkappa n. \tag{5}$$

由假设知确有一个复单位向量 $\boldsymbol{X}$ 使 $\boldsymbol{A}\boldsymbol{X}=\lambda\boldsymbol{X}$，故 $\lambda=\overline{\boldsymbol{X}}'\boldsymbol{A}\boldsymbol{X}$，因之利用(5)式得知 $|\lambda|=|\overline{\boldsymbol{X}}'\boldsymbol{A}\boldsymbol{X}|\leqslant\rho n$.

又因 $\bar{\lambda}=\overline{\boldsymbol{X}}'\overline{\boldsymbol{A}}'\boldsymbol{X}$，故

$$2\cdot R(\lambda)=\lambda+\bar{\lambda}=\overline{\boldsymbol{X}}'(\boldsymbol{A}+\boldsymbol{A}')\boldsymbol{X},$$

$$2i\cdot I(\lambda)=\lambda-\bar{\lambda}=\overline{\boldsymbol{X}}'(\boldsymbol{A}-\overline{\boldsymbol{A}}')\boldsymbol{X},$$

于是再利用(5)式就得到 $|R(\lambda)|\leqslant\sigma n$ 及 $|I(\lambda)|\leqslant\tau n$. 证完.

注意：当 $\boldsymbol{A}$ 是厄米特矩阵时，$a_{rs}=\bar{a}_{sr}$，则 $\tau=0$，故 $I(\lambda)=0$，即 λ 为实数. 当 $\boldsymbol{A}$ 是反厄米特矩阵时，$a_{rs}=-\bar{a}_{sr}$，于是 $\sigma=0$，有 $R(\lambda)=0$，故 λ 为纯虚数. 这从另一个角度又

证得我们过去所熟知的结论.

关于复矩阵的特征值，还有下面的：

定理8 设 $\boldsymbol{A}=(a_{rs})$ 是 n 级复矩阵，

$$\rho_k=\sum_{\substack{s=1\\s\neq k}}^{n}|a_{ks}|,\quad (k=1,\ 2,\ \cdots,\ n),$$

则 $\boldsymbol{A}$ 之每个特征值 λ 至少在由 n 个不等式

$$|z-a_{kk}|\leqslant\rho_k\quad (k=1,\ 2,\ \cdots,\ n)$$

所决定的 n 个圆中的某一个的里面或上面.

证明 令 $\boldsymbol{X}=\begin{pmatrix}x_1\\ \vdots\\ x_n\end{pmatrix}$ 为对应于 λ 的 $\boldsymbol{A}$ 之任一个特征向量，则 $\boldsymbol{AX}=\lambda\boldsymbol{X}$，故

$$(\lambda-a_{rr})x_r=\sum_{\substack{s=1\\s\neq r}}^{n}a_{rs}x_s,\quad (r=1,\ 2,\ \cdots,\ n).\tag{6}$$

今令 $|x_k|=\max(|x_1|,\ |x_2|,\ \cdots,\ |x_n|)$，则 $x_k\neq0$，故从(6)式得知

$$|\lambda-a_{kk}|\cdot|x_k|=\left|\sum_{\substack{s=1\\s\neq k}}^{n}a_{ks}x_s\right|\leqslant\sum_{\substack{s=1\\s\neq k}}^{n}|a_{ks}|\cdot|x_k|$$

$$=|x_k|\cdot\sum_{\substack{s=1\\s\neq k}}^{n}|a_{ks}|=\rho_k|x_k|,$$

因之，
$$|\lambda-a_{kk}|\leqslant\rho_k,$$
证完.

推论1 若 n 级矩阵 $\boldsymbol{A}=(a_{rs})$ 之主对角线上每个元素的模大于它所在的行上其他各元素的模之和，即

$$|a_{kk}|>\sum_{\substack{s=1\\s\neq k}}^{n}|a_{ks}|\quad (k=1,\ 2,\ \cdots,\ n),$$

则 $\boldsymbol{A}$ 是满秩的，并且每个特征值 λ 至少适合 n 个不等式

$$|\lambda-a_{kk}|<|a_{kk}| \quad (k=1, 2, \cdots, n)$$

中的一个.

证明 这时，$\rho_k<|a_{kk}|(k=1, 2, \cdots, n)$，故由定理3即知$n$个不等式

$$|\lambda-a_{kk}|<|a_{kk}|$$

中至少有一个成立，即证得了推论1的后半.

既然有一k使$|\lambda-a_{kk}|<|a_{kk}|$，故$\lambda\neq 0$，这表示了$\boldsymbol{A}$的特征值无一为零，因而$\boldsymbol{A}$是满秩的. 证完.

推论2 设$\boldsymbol{A}=(a_{rs})$是n级实矩阵，且$a_{kk}>\sum\limits_{\substack{s=1\\s\neq k}}^{n}|a_{ks}|$，$(k=1, 2, \cdots, n)$，则$\boldsymbol{A}$之每特征值都在复平面的右半部.

这只要证明每特征值λ之实部$R(\lambda)$是正数就行了. 今令$\lambda=a+bi$，$a=R(\lambda)$，$b=I(\lambda)$. 据推论1，至少有一个r，使

$$|\lambda-a_{rr}|<a_{rr},$$

即
$$|(a-a_{rr})+bi|<a_{rr},$$

故
$$(a-a_{rr})^2+b^2<a_{rr}^2,$$

因而
$$a^2+b^2<2aa_{rr},$$

即有$aa_{rr}>0$. 但$a_{rr}>0$，所以$a=R(\lambda)>0$，证完.

练习四十二

1. 设$\lambda_1, \lambda_2, \cdots, \lambda_n$是矩阵$\boldsymbol{A}=(a_{ij})$之$n$个特征值，试证:

i) $\sum\limits_{r=1}^{n}|\lambda_r|^2\leqslant\sum\limits_{r,s=1}^{n}|a_{rs}|^2$;

ii) $\sum\limits_{r=1}^{n}|R(\lambda_r)|^2\leqslant\sum\limits_{r,s=1}^{n}\left|\dfrac{a_{rs}+\bar{a}_{sr}}{2}\right|^2$;

iii) $\sum\limits_{r=1}^{n}|I(\lambda_r)|^2\leqslant\sum\limits_{r,s=1}^{n}\left|\dfrac{a_{rs}-\bar{a}_{sr}}{2}\right|^2$.

2. 满足定理 3 之推论 1 之条件的矩阵 $\boldsymbol{A}$ 不仅是满秩的，且有 $\det\boldsymbol{A}$ 的模$\geqslant d_1d_2\cdots d_n$，但 $d_k=|a_{kk}|-\sum\limits_{\substack{s=1\\s\neq k}}^{n}|a_{ks}|$，$(k=1, 2, \cdots, n)$。

第 4 节　矩阵的极分解

我们知道：任一个非零的复数 α 恒可写为 $\alpha=\rho\cdot(\cos\theta+i\sin\theta)$形，式中 $\rho(>0)$ 为 α 之极径(或叫做模)，θ 为 α 之幅角；且这样的表写法是唯一的(除幅角相差 2π 的一个倍数外)．这就是所谓数之极分解．注意 $\rho>0$ 表明 ρ 可视为一级的正定矩阵，而 $\cos\theta+i\sin\theta$ 得视为一级的酉矩阵．极分解的概念可以推广到矩阵上去．说具体些，任一满秩矩阵 $\boldsymbol{A}$ 恒可唯一地表写为 $\boldsymbol{A}=\boldsymbol{HU}$ 形，$\boldsymbol{H}$ 与 $\boldsymbol{U}$ 分别是正定厄米特矩阵与酉矩阵．为此，先叙述一些预备知识．

引理 1　设 $\boldsymbol{A}$、$\boldsymbol{B}$ 是两个正定厄米特矩阵(或正定实对称矩阵)，且 $\boldsymbol{A}^m=\boldsymbol{B}^m$($m$ 是正整数)，则 $\boldsymbol{A}=\boldsymbol{B}$.

证明　$\boldsymbol{A}$、$\boldsymbol{B}$ 的特征值都是正数，故令 $\boldsymbol{A}$ 之 n 个正特征值依从小到大的次序来排而为 $\lambda_1\leqslant\lambda_2\leqslant\cdots\leqslant\lambda_n$，同样 $\boldsymbol{B}$ 之 n 个正特征值为 $\mu_1\leqslant\mu_2\leqslant\cdots\leqslant\mu_n$. 由于 $\boldsymbol{A}^m$ 之特征值为 $\lambda_1^m\leqslant\lambda_2^m\leqslant\cdots\leqslant\lambda_n^m$，$\boldsymbol{B}^m$ 之特征值为 $\mu_1^m\leqslant\mu_2^m\leqslant\cdots\leqslant\mu_n^m$，故从 $\boldsymbol{A}^m=\boldsymbol{B}^m$ 就有 $\lambda_i^m=\mu_i^m$ 从而 $\lambda_i=\mu_i(i=1, 2, \cdots, n)$.

于是，存在适当的酉矩阵 $\boldsymbol{U}$、$\boldsymbol{U}_1$，使得

$$\boldsymbol{U}^{-1}\boldsymbol{AU}=\boldsymbol{C},\quad \boldsymbol{U}_1^{-1}\boldsymbol{BU}_1=\boldsymbol{C},$$

$$\boldsymbol{C}=\begin{pmatrix}\lambda^{(1)}\boldsymbol{E}^{(1)} & & 0\\ & \ddots & \\ 0 & & \lambda^{(t)}\boldsymbol{E}^{(t)}\end{pmatrix},$$

$$\lambda^{(i)}\neq\lambda^{(j)}\quad(i\neq j),\quad(\boldsymbol{E}^{(i)}\text{ 为单位阵}).$$

即 $\lambda_1, \cdots, \lambda_n$ 中互异的令为 $\lambda^{(1)}, \cdots, \lambda^{(t)}$. (注意：当 $\boldsymbol{A}$, $\boldsymbol{B}$ 为

实对称时，可选 U 及 U_1 都是实直交的.）

由 $A=UCU^{-1}$, $B=U_1CU_1^{-1}$, 得

$$UC^mU^{-1}=A^m=B^m=U_1C^mU_1^{-1},$$

故

$$U_1^{-1}U\cdot C^m=C^m\cdot U_1^{-1}U,$$

即 $U_1^{-1}U$ 与 $C^m=\begin{pmatrix}(\lambda_1^{(1)})^mE^{(1)} & & \\ & \ddots & \\ & & (\lambda^{(t)})^mE^{(t)}\end{pmatrix}$ 相乘可交换，故从 $(\lambda^{(i)})^m\neq(\lambda^{(j)})^m$（$i\neq j$ 时）就知道

$$U_1^{-1}U=\begin{pmatrix}D_1 & & 0\\ & \ddots & \\ 0 & & D_t\end{pmatrix}$$

形之矩阵，D_j 与 $E^{(j)}$ 同级，因而当然也必有

$$U_1^{-1}U\cdot C=C\cdot U_1^{-1}U,$$

即 $UCU^{-1}=U_1CU_1^{-1}$, $A=B$, 证完.

引理 2 设 A 为正定厄米特矩阵（或正定实对称矩阵），则有唯一个正定厄米特矩阵（或正定实对称矩阵）X, 使 $X^m=A$（m 为正整数）.

证明 有酉矩阵 U 使

$$U^{-1}AU=\mathrm{diag}\{\lambda_1,\ \cdots,\ \lambda_n\},\quad \lambda_i>0,$$

并令 $X=U\cdot\mathrm{diag}\{\sqrt[m]{\lambda_1},\ \cdots,\ \sqrt[m]{\lambda_n}\}\cdot U^{-1}$, 则

$$X^m=U\cdot\mathrm{diag}\{\lambda_1,\ \cdots,\ \lambda_n\}\cdot U^{-1},$$

$$U^{-1}X^mU=\mathrm{diag}\{\lambda_1,\ \cdots,\ \lambda_n\}=U^{-1}AU,$$

故 $X^m=A$, 又 X 显然是正定厄米特矩阵（当 A 为正定实对称矩阵时，可选 U 为实直交矩阵，因而 X 为正定实对称的）. 至于 X 之唯一性，由引理 1 即得. 证完.

引理 3 厄米特矩阵 H 为正定或半正定的充要条件是有一满秩矩阵或降秩矩阵 A, 使 $H=\overline{A}'A$（或 $H=A\overline{A}'$）

证明 先证充分性. 令

$$X=\begin{pmatrix}x_1\\ \vdots\\ x_n\end{pmatrix},\quad Y=\begin{pmatrix}y_1\\ \vdots\\ y_n\end{pmatrix},$$

并令 $Y=AX$, 则

$$\overline{X}'HX=\overline{X}'\overline{A}'AX=\overline{Y}'Y. \tag{7}$$

然当 A 为满秩时, $Y=0$ 是在且仅在 $X=0$ 时, 故由(7)即知 $X\neq 0$ 恒使 $\overline{X}'HX>0$, 因而 H 为正定的. 但在 A 为降秩时, 有某 $X\neq 0$ 却使 $Y=0$, 因之据(7)式有某 $X\neq 0$ 使 $\overline{X}'HX=0$ 而一般恒有 $\overline{X}'HX\geqslant 0$, 说明了 H 为半正定的.

再证必要性. 先设 H 为正定的, 则有满秩阵 Q 使

$$\overline{Q}'HQ=E=\operatorname{diag}\{1,\cdots,1\},$$

故

$$H=\overline{(Q^{-1})}'Q^{-1}=\overline{A}'A(A=Q^{-1}).$$

再设 H 是半正定的, 于是有满秩矩阵 $Q=(q_{ij})$ 使

$$H=\overline{Q}'\cdot\operatorname{diag}\{\underbrace{1,\cdots,1}_{r\text{个}},0,\cdots,0\}\cdot Q,\quad r<n,$$

因之,

$$H=\begin{pmatrix}\bar q_{11}\cdots\bar q_{r1} & \bar q_{r+1,1}\cdots\bar q_{n1}\\ \vdots\quad\ \vdots & \vdots\quad\ \vdots\\ \bar q_{1r}\cdots\bar q_{rr} & \bar q_{r+1,r}\cdots\bar q_{nr}\\ \vdots\quad\ \vdots & \vdots\quad\ \vdots\\ \bar q_{1n}\cdots\bar q_{rn} & \bar q_{r+1,n}\cdots\bar q_{nn}\end{pmatrix}\begin{pmatrix}q_{11}\cdots q_{1r} & q_{1,r+1}\cdots q_{1n}\\ \vdots\quad\ \vdots & \vdots\quad\ \vdots\\ q_{r1}\cdots q_{rr} & q_{r,r+1}\cdots q_{rn}\\ 0\ \cdots\ 0 & 0\ \cdots\ 0\\ \vdots\quad\ \vdots & \vdots\quad\ \vdots\\ 0\ \cdots\ 0 & 0\ \cdots\ 0\end{pmatrix}$$

$$=\begin{pmatrix}\bar q_{11}\cdots\bar q_{r1} & 0\cdots 0\\ \vdots\quad\ \vdots & \vdots\quad\ \vdots\\ \bar q_{1r}\cdots\bar q_{rr} & 0\cdots 0\\ \vdots\quad\ \vdots & \vdots\quad\ \vdots\\ \bar q_{1n}\cdots\bar q_{rn} & 0\cdots 0\end{pmatrix}\begin{pmatrix}q_{11}\cdots q_{1r}\cdots q_{1n}\\ \vdots\quad\ \vdots\quad\ \vdots\\ q_{r1}\cdots q_{rr}\cdots q_{rn}\\ 0\ \cdots\ 0\ \cdots\ 0\\ \vdots\quad\ \vdots\quad\ \vdots\\ 0\ \cdots\ 0\ \cdots\ 0\end{pmatrix}=\overline{A}'A,$$

但
$$A=\begin{pmatrix} q_{11}\cdots\cdots q_{1n} \\ \cdots\cdots\cdots \\ q_{r1}\cdots\cdots q_{rn} \\ 0\cdots\cdots\cdots 0 \\ \cdots\cdots\cdots \\ 0\cdots\cdots\cdots 0 \end{pmatrix}$$
显为降秩的. 证完.

推论 实对称矩阵 S 为正定或半正定的充要条件是有实满秩或实降秩矩阵 A 使 $S=A'A$(或 $=AA'$).

现在可以解决矩阵之极分解问题, 即:

定理 1 (矩阵的极分解). 任何满秩矩阵 A 恒可表为

i) $A=HU$, H 是正定厄米特的, U 是酉矩阵;

或形

ii) $A=U_1H_1$, H_1 是正定厄米特的, U_1 是酉矩阵;

且 i) 或 ii) 之表示法都是唯一的; 又, i) 与 ii) 中的 $U=U_1$.

证明 由引理 3 知 $K=\bar{A}'A$ 是正定的, 故由引理 2 知有正定厄米特矩阵 H_1, 使 $\bar{A}'A=K=H_1^2$; 再令 $U_1=AH_1^{-1}$, 则
$$\begin{aligned}\bar{U}_1'U_1&=(\overline{H_1^{-1}})'\,A'AH_1^{-1}=(\bar{H}_1')^{-1}H_1^2H_1^{-1}\\&=H_1^{-1}H_1^2H_1^{-1}=E,\end{aligned}$$
说明了 U_1 为酉矩阵, 故 $A=U_1H_1$, 即证明了 ii).

同理, 根据引理 3, $A\bar{A}'$ 是正定的; 根据引理 2, 又有正定的 H 使得 $A\bar{A}'=H^2$. 再令 $U=H^{-1}A$, 则
$$\begin{aligned}\bar{U}'U&=\bar{A}'(\bar{H}')^{-1}H^{-1}A=\bar{A}'H^{-2}A=\bar{A}'(A\bar{A}')^{-1}\cdot A\\&=\bar{A}'\cdot(\bar{A}')^{-1}\cdot A^{-1}A=E,\end{aligned}$$
表示了 U 是酉矩阵, 故 $A=HU$, 即证明了 i).

又, 若 $A=HU=H_0U_0$(H_0 为正定厄米特的, U_0 为酉矩阵), 则 $H_0=HUU_0^{-1}=HV$, $V=UU_0^{-1}$ 为酉矩阵, 于是有

$$H_0^2=H_0\overline{H}_0'=HV\overline{V}'\overline{H}'=HVV^{-1}H=HEH=H^2,$$

由引理 1, 有 $H_0=H$, 因之 $U_0=U$. 即证明了 i)之表示的唯一性. 同理可知 ii)之表示的唯一性.

最后, 从 i)有$A=HU=U\cdot U^{-1}HU=UH_2, H_2=U^{-1}HU$为正定的, 故由 ii)之唯一性, 就有 $U=U_1(H_1=U^{-1}HU)$, 定理 1 证完.

同样, 又有:

定理 2 任何实满秩矩阵 A 恒可唯一地表写为 $A=ST$ (S 为正定实对称, T 为实直交矩阵), 又可唯一地表写为 $A=T_1\cdot S_1$(T_1 为实直交, S_1 为正定实对称矩阵), 且 $T=T_1$.

证明留给读者当练习题.

关于降秩矩阵之极分解又怎样呢?

今设 A 是降秩(n 级)的. 于是, 由引理 3, $A\overline{A}'$ 为半正定, 故有半正定的 H 使 $A\overline{A}'=H^2$(与引理 2 之证法完全一样, 且 H 还是唯一的). 由于有酉矩阵 V 使

$$VHV^{-1}=\mathrm{diag}\{\rho_1, \cdots, \rho_r, 0, \cdots, 0\}, \quad \rho_i>0 \text{ 且 } r<n,$$

故若令
$$VAV^{-1}=\begin{pmatrix} c_{11} & \cdots & c_{1n} \\ \vdots & & \vdots \\ c_{n1} & \cdots & c_{nn} \end{pmatrix},$$

则

$$\begin{pmatrix} \sum c_{1i}\bar{c}_{1i} & \sum c_{1i}\bar{c}_{2i} & \cdots & \sum c_{1i}\bar{c}_{ni} \\ \sum c_{2i}\bar{c}_{1i} & \sum c_{2i}\bar{c}_{2i} & \cdots & \sum c_{2i}\bar{c}_{ni} \\ \vdots & \vdots & & \vdots \\ \sum c_{ni}\bar{c}_{1i} & \sum c_{ni}\bar{c}_{2i} & \cdots & \sum c_{ni}\bar{c}_{ni} \end{pmatrix}$$

$$=\begin{pmatrix} c_{11} & c_{12} & \cdots & c_{1n} \\ c_{21} & c_{22} & \cdots & c_{2n} \\ \vdots & \vdots & & \vdots \\ c_{n1} & c_{n2} & \cdots & c_{nn} \end{pmatrix}\begin{pmatrix} \bar{c}_{11} & \bar{c}_{21} & \cdots & \bar{c}_{n1} \\ \bar{c}_{12} & \bar{c}_{22} & \cdots & \bar{c}_{n2} \\ \vdots & \vdots & & \vdots \\ \bar{c}_{1n} & \bar{c}_{2n} & \cdots & \bar{c}_{nn} \end{pmatrix}$$

$$=(\boldsymbol{V}\boldsymbol{A}\overline{\boldsymbol{V}}')\cdot\overline{(\boldsymbol{V}\boldsymbol{A}\overline{\boldsymbol{V}}')}'=\boldsymbol{V}\boldsymbol{A}\overline{\boldsymbol{A}}'\boldsymbol{V}^{-1}=\boldsymbol{V}\boldsymbol{H}^2\boldsymbol{V}^{-1}$$

$$=(\boldsymbol{V}\boldsymbol{H}\boldsymbol{V}^{-1})^2=\operatorname{diag}\{\rho_1^2,\ \cdots,\ \rho_r^2,\ 0,\ \cdots,\ 0\},$$

故
$$\sum_{i=1}^{n}c_{ti}\bar{c}_{ti}=\rho_t^2\quad(t=1,\ \cdots,\ r),$$

$$\sum_{i=1}^{n}c_{si}\bar{c}_{si}=0\quad(s=r+1,\ \cdots,\ n),$$

$$\sum_{i=1}^{n}c_{\lambda i}\bar{c}_{\mu i}=0\quad(\lambda\neq\mu).$$

由
$$\sum_{i=1}^{n}c_{si}\bar{c}_{si}=0\quad(s=r+1,\ \cdots,\ n)$$

得每
$$c_{si}=0\quad(s=r+1,\ \cdots,\ n);$$

再令
$$c_{ti}/\rho_t=d_{ti}\quad(t=1,\ \cdots,\ r;\ i=1,\ \cdots,\ n),$$

就得到

$$\operatorname{diag}\{\rho_1^{-1},\ \cdots,\ \rho_r^{-1},\ 1,\ \cdots,\ 1\}\cdot\boldsymbol{V}\boldsymbol{A}\boldsymbol{V}^{-1}$$

$$=\begin{pmatrix} d_{11} & d_{12} & \cdots & d_{1n}\\ \vdots & \vdots & & \vdots\\ d_{r1} & d_{r2} & \cdots & d_{rn}\\ 0 & 0 & \cdots & 0\\ \vdots & \vdots & & \vdots\\ 0 & 0 & \cdots & 0\end{pmatrix},$$

其中前 r 个行满足关系式

$$\sum_{i=1}^{n}d_{ti}\bar{d}_{ti}=1\quad(t=1,\ \cdots,\ r)$$

及
$$\sum_{i=1}^{n}d_{\lambda i}\bar{d}_{\mu i}=0\quad(\lambda,\ \mu=1,\ \cdots,\ r\text{ 且 }\lambda\neq\mu).$$

因而，以这 r 个行为一酉矩阵 $\boldsymbol{W}=(w_{ij})$ 的前 r 个行作酉矩阵 $\boldsymbol{W}$（这样的 $\boldsymbol{W}$ 当然不是唯一的），且由于

$$VAV^{-1}=\mathrm{diag}\{\rho_1,\cdots,\rho_r,1,\cdots,1\}\cdot\begin{pmatrix}d_{11}&d_{12}&\cdots&d_{1n}\\ \vdots&\vdots&&\vdots\\ d_{r1}&d_{r2}&\cdots&d_{rn}\\ 0&0&\cdots&0\\ \vdots&\vdots&&\vdots\\ 0&0&\cdots&0\end{pmatrix}$$

$$=\mathrm{diag}\{\rho_1,\cdots,\rho_r,0,\cdots,0\}$$

$$\cdot\begin{pmatrix}d_{11}&d_{12}&\cdots&d_{1n}\\ \vdots&\vdots&&\vdots\\ d_{r1}&d_{r2}&\cdots&d_{rn}\\ w_{r+1,1}&w_{r+1,2}&\cdots&w_{r+1,n}\\ \vdots&\vdots&&\vdots\\ w_{n1}&w_{n2}&\cdots&w_{nn}\end{pmatrix}$$

$$\mathrm{diag}\{\rho_1,\cdots,\rho_r,0,\cdots,0\}\cdot\boldsymbol{W}=\boldsymbol{VHV}^{-1}\boldsymbol{W},$$

又有 $\boldsymbol{A}=\boldsymbol{H}\cdot\boldsymbol{V}^{-1}\boldsymbol{WV}=\boldsymbol{HU}$($\boldsymbol{U}=\boldsymbol{V}^{-1}\boldsymbol{WV}$ 为酉矩阵).

但 $\boldsymbol{A}=\boldsymbol{HU}=\boldsymbol{U}\cdot\boldsymbol{U}^{-1}\boldsymbol{HU}=\boldsymbol{U}\cdot\boldsymbol{H}_1$, $\boldsymbol{H}_1=\boldsymbol{U}^{-1}\boldsymbol{HU}$ 也是半正定的. 这就证明了降秩矩阵 $\boldsymbol{A}$ 恒可表为

$$\boldsymbol{A}=\boldsymbol{HU}\quad 或\quad \boldsymbol{A}=\boldsymbol{U}_1\boldsymbol{H}_1,$$

但 $\boldsymbol{H}$ 与 $\boldsymbol{H}_1$ 都是半正定厄米特矩阵, $\boldsymbol{U}$ 与 $\boldsymbol{U}_1$ 都是酉矩阵.

又, 若 $\boldsymbol{A}=\boldsymbol{HU}=\boldsymbol{H}_0\boldsymbol{U}_0$ ($\boldsymbol{H}$ 与 $\boldsymbol{H}_0$ 皆为半正定的厄米特矩阵, $\boldsymbol{U}$ 与 $\boldsymbol{U}_0$ 都是酉矩阵), 则 $\boldsymbol{H}_0=\boldsymbol{HR}$($\boldsymbol{R}=\boldsymbol{UU}_0^{-1}$ 为酉矩阵), 故 $\boldsymbol{H}_0^2=\boldsymbol{H}_0\overline{\boldsymbol{H}}'_0=\boldsymbol{HR}\overline{\boldsymbol{R}}'\overline{\boldsymbol{H}}'=\boldsymbol{H}\overline{\boldsymbol{H}}'=\boldsymbol{H}^2$; 再与证引理 1 完全一样的方法, 可知 $\boldsymbol{H}_0=\boldsymbol{H}$, 这说明了分解 $\boldsymbol{A}=\boldsymbol{HU}$ 中 $\boldsymbol{H}$ 是唯一确定的, 而 $\boldsymbol{U}$ 则不然.

故得知有:

定理 3 任何降秩矩阵 $\boldsymbol{A}$ 恒可表写为 $\boldsymbol{A}=\boldsymbol{HU}$ 形, 或可表写为 $\boldsymbol{A}=\boldsymbol{U}_1\boldsymbol{H}_1$ 形($\boldsymbol{H}$ 与 $\boldsymbol{H}_1$ 是半正定的, $\boldsymbol{U}$ 与 $\boldsymbol{U}_1$ 为酉矩阵), 且 $\boldsymbol{H}$ 与 $\boldsymbol{H}_1$ 皆唯一确定, 但 $\boldsymbol{U}$ 与 $\boldsymbol{U}_1$ 则不然.

对实矩阵，又有：

定理4 任何降秩实矩阵 $\boldsymbol{A}$ 恒可表写为 $\boldsymbol{A}=\boldsymbol{S}\boldsymbol{T}$ 形或 $\boldsymbol{A}=\boldsymbol{T}_1\boldsymbol{S}_1$ 形，$\boldsymbol{S}$ 与 $\boldsymbol{S}_1$ 都是半正定实对称矩阵，而 $\boldsymbol{T}$ 与 $\boldsymbol{T}_1$ 都是实直交矩阵．且 $\boldsymbol{S}$ 与 $\boldsymbol{S}_1$ 都唯一确定，而 $\boldsymbol{T}$ 与 $\boldsymbol{T}_1$ 则不然．

由矩阵之极分解，可看出正定(半正定)厄米特矩阵与酉矩阵在矩阵中的地位．因此，下面再对正定厄米特矩阵与酉矩阵自身的属性作进一步的探索(当然，厄米特矩阵与酉矩阵的一些基本性质，已在第六、七、八章里陆续地叙述过了)．

在下面的讨论中，我们要用到 Hölder 不等式：设 α_1, α_2, …, α_n 与 β_1, β_2, …, β_n 是两组 n 个正数，则有：

$$(\alpha_1\alpha_2\cdots\alpha_n)^{1/n}+(\beta_1\beta_2\cdots\beta_n)^{1/n}$$
$$\leqslant\{(\alpha_1+\beta_1)(\alpha_2+\beta_2)\cdots(\alpha_n+\beta_n)\}^{1/n},$$

且等号“$=$”当 $\beta_i=\varkappa\alpha_i(i=1, \cdots, n)$ 时成立．

证明 用数学归纳法来证明．$n=1$ 时显然成立．再设 $n=2$，来证明：

$$(\alpha_1\alpha_2)^{\frac{1}{2}}+(\beta_1\beta_2)^{\frac{1}{2}}\leqslant\{(\alpha_1+\beta_1)(\alpha_2+\beta_2)\}^{\frac{1}{2}}.$$

上式左、右两端平方，得

$$\alpha_1\alpha_2+\beta_1\beta_2+2(\alpha_1\alpha_2\beta_1\beta_2)^{\frac{1}{2}}\leqslant(\alpha_1+\beta_1)(\alpha_2+\beta_2),$$

即
$$2(\alpha_1\alpha_2\beta_1\beta_2)^{\frac{1}{2}}\leqslant\alpha_1\beta_2+\alpha_2\beta_1,$$

因之
$$[(\alpha_1\beta_2)^{\frac{1}{2}}-(\alpha_2\beta_1)^{\frac{1}{2}}]^2\geqslant 0,$$

而这是显然正确的，说明 $n=2$ 时 Hölder 不等式也成立；且等号当且仅当 $\alpha_1\beta_2=\alpha_2\beta_1$ 即 $\beta_i=\varkappa\alpha_i(i=1, 2)$ 时成立．再设 n 时成立，而利用这归纳假定，就有

$$\{(\alpha_1+\beta_1)(\alpha_2+\beta_2)\cdots(\alpha_n+\beta_n)(\alpha_{n+1}+\beta_{n+1})\}^{\frac{1}{n+1}}$$
$$=\{(\alpha_1+\beta_1)\cdots(\alpha_n+\beta_n)\}^{\frac{1}{n}\cdot\frac{n}{n+1}}\cdot(\alpha_{n+1}+\beta_{n+1})^{\frac{1}{n+1}}$$

$$\geqslant [(\alpha_1\alpha_2\cdots\alpha_n)^{\frac{1}{n}}+(\beta_1\beta_2\cdots\beta_n)^{\frac{1}{n}}]^{\frac{n}{n+1}}\cdot(\alpha_{n+1}+\beta_{n+1})^{\frac{1}{n+1}}$$

$$=\{[(\alpha_1\alpha_2\cdots\alpha_n)^{\frac{1}{n}}+(\beta_1\beta_2\cdots\beta_n)^{\frac{1}{n}}]^n\}^{\frac{1}{n+1}}\cdot(\alpha_{n+1}+\beta_{n+1})^{\frac{1}{n+1}}$$

$$\geqslant(\alpha_1\alpha_2\cdots\alpha_n+\beta_1\beta_2\cdots\beta_n)^{\frac{1}{n+1}}(\alpha_{n+1}+\beta_{n+1})^{\frac{1}{n+1}}$$

$$=\{(\alpha_1\alpha_2\cdots\alpha_n+\beta_1\beta_2\cdots\beta_n)(\alpha_{n+1}+\beta_{n+1})\}^{\frac{1}{n+1}}$$

$$=\{[(\alpha_1\alpha_2\cdots\alpha_n+\beta_1\beta_2\cdots\beta_n)(\alpha_{n+1}+\beta_{n+1})]^{\frac{1}{2}}\}^{\frac{2}{n+1}}$$

$$\geqslant\{(\alpha_1\alpha_2\cdots\alpha_n\alpha_{n+1})^{\frac{1}{2}}+(\beta_1\beta_2\cdots\beta_n\beta_{n+1})^{\frac{1}{2}}\}^{\frac{2}{n+1}}$$

$$\geqslant\{[(\alpha_1\alpha_2\cdots\alpha_n\alpha_{n+1})^{\frac{1}{2}}+(\beta_1\beta_2\cdots\beta_n\beta_{n+1})^{\frac{1}{2}}]^2\}^{\frac{1}{n+1}}$$

$$\geqslant\{\alpha_1\alpha_2\cdots\alpha_n\alpha_{n+1}+\beta_1\beta_2\cdots\beta_n\beta_{n+1}\}^{\frac{1}{n+1}},$$

即证明了不等式在 n 时成立的归纳假定下在 $n+1$ 时也必成立. 又,明显地可知不等式中“等号”当 $\beta_i=\varkappa\alpha_i$ 时成立.

下面研究正定厄米特矩阵,为此,先证:

引理4 设 $\boldsymbol{A}$、$\boldsymbol{B}$ 为厄米特矩阵,且 $\boldsymbol{B}$ 为正定的,则存在行列式等于1的矩阵 $\boldsymbol{P}$,使得

$$\overline{\boldsymbol{P}}'\boldsymbol{A}\boldsymbol{P}=\begin{pmatrix}\lambda_1 & & \\ & \ddots & \\ & & \lambda_n\end{pmatrix} \quad 及 \quad \overline{\boldsymbol{P}}'\boldsymbol{B}\boldsymbol{P}=\begin{pmatrix}\mu_1 & & \\ & \ddots & \\ & & \mu_n\end{pmatrix}$$

(因之 λ_i 皆为实数, μ_i 皆为正数).

证明 据第八章第4节的定理,知有一满秩矩阵 $\boldsymbol{Q}$ 使 $\overline{\boldsymbol{Q}}'\boldsymbol{A}\boldsymbol{Q}$ 为对角阵,且 $\overline{\boldsymbol{Q}}'\boldsymbol{B}\boldsymbol{Q}=\boldsymbol{E}$. 再令

$$\boldsymbol{P}=(1/\det\boldsymbol{Q})^{1/n}\cdot\boldsymbol{Q} \quad (n 为 \boldsymbol{Q} 之级数),$$

则显有
$$\det\boldsymbol{P}=\left[\left(\frac{1}{\det\boldsymbol{Q}}\right)^{1/n}\right]^n\cdot\det\boldsymbol{Q}=1,$$

且 $\overline{\boldsymbol{P}}'\boldsymbol{A}\boldsymbol{P}$ 是对角阵,及

$$\overline{P}'BP=\begin{pmatrix}\mu_1 & & \\ & \ddots & \\ & & \mu_n\end{pmatrix}.$$

实际上，$\overline{P}'BP=cE$ 为纯量矩阵（$c>0$）.

现在就来解决关于正定厄米特矩阵的一些重要结果，即有下列定理：

定理 5 设 A、B 为 n 级正定厄米特矩阵，则

$$|A|^{1/n}+|B|^{1/n}\leqslant|A+B|^{1/n},$$

且等号在 $B=\varkappa A$ 时成立.

证明 由引理 4，知有满秩矩阵 P，使得

$$|P|=1,\quad \overline{P}'AP=\begin{pmatrix}\alpha_1 & & 0\\ & \ddots & \\ 0 & & \alpha_n\end{pmatrix},$$

$$\overline{P}'BP=\begin{pmatrix}\beta_1 & & 0\\ & \ddots & \\ 0 & & \beta_n\end{pmatrix},$$

这时 α_i 与 β_i 都是正数．由是就有

$$|A|=\alpha_1\alpha_2\cdots\alpha_n,\qquad |B|=\beta_1\beta_2\cdots\beta_n,$$

而

$$\begin{aligned}|A+B|&=|\overline{P}'(A+B)P|=|\overline{P}'AP+\overline{P}'BP|\\&=(\alpha_1+\beta_1)(\alpha_2+\beta_2)\cdots(\alpha_n+\beta_n),\end{aligned}$$

于是据 Hölder 不等式，即证得定理 5.

定理 6 设 A、B 为 n 级正定厄米特矩阵（$n>1$），则 $|A|+|B|<|A+B|$.

证明 由定理 5，知 $|A|^{1/n}+|B|^{1/n}\leqslant|A+B|^{1/n}$．由于 $|A|$ 及 $|B|$ 都是正数，故将上式两端 n 乘幂后，即得

$$|A|+|B|<|A+B|.$$

定理 7 设 A、B 为两个正定（同级）厄米特矩阵，α 为正

实数，适合 $0<\alpha<1$，则

$$|\alpha \boldsymbol{A}+(1-\alpha)\boldsymbol{B}|\geqslant|\boldsymbol{A}|^{\alpha}\cdot|\boldsymbol{B}|^{1-\alpha},$$

且等号仅限于 $\boldsymbol{A}=\boldsymbol{B}$ 时才成立.

证明 注意 $\alpha\boldsymbol{A}$ 与 $(1-\alpha)\boldsymbol{B}$ 也都是正定厄米特矩阵，于是利用定理5中 $\boldsymbol{P}$ 之存在性，就有

$$|\alpha \boldsymbol{A}+(1-\alpha)\boldsymbol{B}|=|\overline{\boldsymbol{P}}'(\alpha\boldsymbol{A}+(1-\alpha)\boldsymbol{B})\boldsymbol{P}|$$
$$=\prod_{i=1}^{n}[\alpha\alpha_i+(1-\alpha)\beta_i];$$

另方面，又有 $\qquad |\boldsymbol{A}|^{\alpha}\cdot|\boldsymbol{B}|^{1-\alpha}=\prod_{i=1}^{n}\alpha_i^{\alpha}\beta_i^{1-\alpha}.$

然 $\alpha\alpha_i+(1-\alpha)\beta_i\geqslant\alpha_i^{\alpha}\beta_i^{1-\alpha}$†，且等号仅在 $\beta_i=\alpha_i$ 时成立，定理7获证.

定理8 设 $\boldsymbol{A}_1,\boldsymbol{A}_2,\cdots,\boldsymbol{A}_k$ 是 k 个 n 级正定厄米特矩阵，则有

$$|\boldsymbol{A}_1+\boldsymbol{A}_2+\cdots+\boldsymbol{A}_k|^{1/n}\geqslant|\boldsymbol{A}_1|^{1/n}+|\boldsymbol{A}_2|^{1/n}+\cdots+|\boldsymbol{A}_k|^{1/n},$$

且等号仅在 $\boldsymbol{A}_i$ 都为某矩阵 $\boldsymbol{B}$ 之倍式时成立.

证明 用数学归纳法来证. $k=2$ 时，由定理5即知其正确性. 今设 $k>2$，并归纳地假定对不超过 $k-1$ 个矩阵时定

---

† 这实质上是算术平均与几何平均之关系. 一般，若 x 与 y 为正数，则

$$(px+qy)/(p+q)\geqslant(x^py^q)^{1/(p+q)}.$$

证明 如令 $\frac{p}{p+q}=\lambda$，则 $\frac{q}{p+q}=1-\lambda$ 且 $0<\lambda<1$，因而问题变成了 $\lambda x+(1-\lambda)y\geqslant x^{\lambda}y^{1-\lambda}$. 如 $x\geqslant y$，再两端除以 y，并令 $t=\frac{x}{y}$，则有 $\lambda t+(1-\lambda)\geqslant t^{\lambda}$，$t\geqslant1$；若 $x\leqslant y$，则将上式两端除以 x，并令 $t=\frac{y}{x}$，而令 $\mu=\frac{q}{p+q}$，因之 $\frac{p}{p+q}=1-\mu$，就变成了 $\mu t+(1-\mu)\geqslant t^{\mu}$，$t\geqslant1$. 总之，我们的问题就变成要证明：若 $0<\lambda<1$，$t\geqslant1$，则恒有

$$\lambda t+(1-\lambda)\geqslant t^{\lambda}.$$

如令 $f(t)=\lambda t+(1-\lambda)-t^{\lambda}$，则 $f'(t)=\lambda-\lambda t^{\lambda-1}=\lambda\left(1-\frac{1}{t^{1-\lambda}}\right)\geqslant0$，故 $f(t)$ 当 $t\geqslant1$ 时为 t 的增函数；但 $f(1)=0$，故 $t\geqslant1$ 时恒有 $f(t)\geqslant f(1)=0$，证毕.

理 8 是成立的. 于是因 $\boldsymbol{A}_1+\boldsymbol{A}_2+\cdots+\boldsymbol{A}_{k-1}$ 亦为正定的, 故由定理 5 及归纳法的假定, 得知

$$
\begin{aligned}
&|\boldsymbol{A}_1+\cdots+\boldsymbol{A}_{k-1}+\boldsymbol{A}_k|^{1/n}\\
&\quad\geqslant|\boldsymbol{A}_1+\cdots+\boldsymbol{A}_{k-1}|^{1/n}+|\boldsymbol{A}_k|^{1/n}\\
&\quad\geqslant|\boldsymbol{A}_1|^{1/n}+\cdots+|\boldsymbol{A}_{k-1}|^{1/n}+|\boldsymbol{A}_k|^{1/n},
\end{aligned}
$$

其中等号限在 $\boldsymbol{A}_k=d(\boldsymbol{A}_1+\cdots+\boldsymbol{A}_{k-1})$, $\boldsymbol{A}_i=c_i\boldsymbol{B}(i=1, \cdots, k-1)$时成立, 因而 $\boldsymbol{A}_i=c_i\boldsymbol{B}(i=1, \cdots, k-1, k)$. 证完.

定理 9 设 $\boldsymbol{A}=(a_{ij})$ 为正定厄米特矩阵$(i, j=1, \cdots, n)$, 且 $1\leqslant r\leqslant n-1$, 则

$$
\begin{vmatrix} a_{11} & \cdots & a_{1r}\\ \vdots & & \vdots\\ a_{r1} & \cdots & a_{rr}\end{vmatrix}\cdot\begin{vmatrix} a_{r+1,r+1} & \cdots & a_{r+1,n}\\ \vdots & & \vdots\\ a_{n,r+1} & \cdots & a_{nn}\end{vmatrix}\geqslant|\boldsymbol{A}|.
$$

证明 令

$$
\boldsymbol{P}=\begin{pmatrix} \overbrace{1 \quad \cdot \quad \cdot \quad \cdot \quad 1}^{r\text{个}} & \\ & \overbrace{-1 \quad \cdot \quad \cdot \quad \cdot \quad -1}^{n-r\text{个}}\end{pmatrix},
$$

则 $\overline{\boldsymbol{P}}=\boldsymbol{P}=\boldsymbol{P}'$, 故

$$
\boldsymbol{B}=\overline{\boldsymbol{P}}'\boldsymbol{A}\boldsymbol{P}=\begin{pmatrix}
a_{11} & \cdots & a_{1r} & -a_{1,r+1} & \cdots & -a_{1n}\\
\vdots & & \vdots & \vdots & & \vdots\\
a_{r1} & \cdots & a_{rr} & -a_{r,r+1} & \cdots & -a_{rn}\\
-a_{r+1,1} & \cdots & -a_{r+1,r} & a_{r+1,r+1} & \cdots & a_{r+1,n}\\
\vdots & & \vdots & \vdots & & \vdots\\
-a_{n1} & \cdots & -a_{nr} & a_{n,r+1} & \cdots & a_{nn}
\end{pmatrix}
$$

为正定的, 且有 $|\boldsymbol{B}|=|\boldsymbol{A}|$. 于是, 若令

$$\Delta=\begin{vmatrix}a_{11}&\cdots&a_{1r}\\ \vdots&&\vdots\\ a_{r1}&\cdots&a_{rr}\end{vmatrix}\cdot\begin{vmatrix}a_{r+1,r+1}&\cdots&a_{r+1,n}\\ \vdots&&\vdots\\ a_{n,r+1}&\cdots&a_{nn}\end{vmatrix},$$

则显有 $|\boldsymbol{A}+\boldsymbol{B}|=2^n\cdot\Delta$, 因而

$$|\boldsymbol{A}+\boldsymbol{B}|^{1/n}=2\cdot\Delta^{1/n}.$$

由定理 5 又有

$$|\boldsymbol{A}+\boldsymbol{B}|^{1/n}\geqslant|\boldsymbol{A}|^{1/n}+|\boldsymbol{B}|^{1/n}=2|\boldsymbol{A}|^{1/n},$$

故 $\Delta^{1/n}\geqslant|\boldsymbol{A}|^{1/n}$, 因之 $\Delta\geqslant|\boldsymbol{A}|$. 证完.

定理 10 设 $\boldsymbol{A}=(a_{ij})$ 与 $\boldsymbol{B}=(b_{ij})$ 为两个 n 级正定厄米特矩阵, 则 $\boldsymbol{A}\circ\boldsymbol{B}=(a_{ij}b_{ij})$ 亦为正定厄米特矩阵. 且当 $\boldsymbol{A}$、$\boldsymbol{B}$ 中有一为半正定时, $\boldsymbol{A}\circ\boldsymbol{B}$ 亦为半正定的. 并当 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 都是正定时又有

$$|\boldsymbol{A}\circ\boldsymbol{B}|\geqslant b_{11}\cdots b_{nn}\cdot|\boldsymbol{A}|\ (\text{或}\ |\boldsymbol{A}\circ\boldsymbol{B}|\geqslant a_{11}a_{22}\cdots a_{nn}\cdot|\boldsymbol{B}|).$$

证明 因 $\boldsymbol{B}$ 正定, 由引理 3, 有 $\boldsymbol{B}=\overline{\boldsymbol{Q}}'\boldsymbol{Q}$, $\boldsymbol{Q}=(q_{ij})$是满秩的. 于是

$$b_{ij}=\sum_{k=1}^{n}\overline{q}_{ki}\cdot q_{kj}.$$

再令

$$\boldsymbol{X}=\begin{pmatrix}x_1\\x_2\\ \vdots\\x_n\end{pmatrix},\quad \boldsymbol{Y}_k=\begin{pmatrix}y_1^{(k)}\\y_2^{(k)}\\ \vdots\\y_n^{(k)}\end{pmatrix}=\begin{pmatrix}q_{k1}x_1\\q_{k2}x_2\\ \vdots\\q_{kn}x_n\end{pmatrix},\quad (k=1,2,\cdots,n),$$

于是当 $\boldsymbol{X}\neq\boldsymbol{0}$ 时, $\boldsymbol{Y}_1$, $\boldsymbol{Y}_2$, $\cdots$, $\boldsymbol{Y}_n$ 中至少有一个不为零(因若不然, 即 $\boldsymbol{Y}_1=\boldsymbol{Y}_2=\cdots=\boldsymbol{Y}_n=\boldsymbol{0}$, 则当然有

$$q_{k1}x_1+q_{k2}x_2+\cdots+q_{kn}x_n=0\quad(k=1,\ \cdots,\ n),$$

即 $\boldsymbol{Q}\boldsymbol{X}=\boldsymbol{0}$, 故由 $\boldsymbol{Q}$ 之满秩性得 $\boldsymbol{X}=\boldsymbol{0}$, 产生了矛盾). 因之, 当 $\boldsymbol{X}\neq\boldsymbol{0}$ 时,

$$\overline{X}'(A\circ B)X=\sum_{i,j=1}^{n}a_{ij}b_{ij}\overline{x}_i x_j=\sum_{i,j=1}^{n}a_{ij}\left(\sum_{k=1}^{n}\overline{q}_{ki}q_{kj}\right)\overline{x}_i x_j$$

$$=\sum_{k=1}^{n}\left(\sum_{i,j=1}^{n}a_{ij}\overline{q_{ki}x_i}\cdot q_{kj}x_j\right)=\sum_{k=1}^{n}\left(\sum_{i,j=1}^{n}a_{ij}\overline{y_i^{(k)}}y_j^{(k)}\right)$$

$$=\sum_{k=1}^{n}\overline{Y}'_kAY_k>0,$$

这就证明了 $A\circ B$ 是正定的.

如果 B(例如)为半正定，则上面的 Q 是降秩的，故虽 $X\neq 0$，而 $QX=0$ 之可能性是有的，亦即 Y_1, …, Y_n 都有可能为零向量，因之当 $X\neq 0$ 时，可能是

$$\overline{X}'(A\circ B)X=\sum_{k=1}^{n}\overline{Y}'_kAY_k=0,$$

这说明了 $A\circ B$ 是半正定的.

于是，剩下只需解决 $|A\circ B|\geq b_{11}b_{22}\cdots b_{nn}|A|$.

现用数学归纳法来证. 当 $n=1$ 时是显然的. 今假定对 $n-1$ 级矩阵时上不等式成立. 于是，令 $C=A\circ B=(c_{ij})$, $c_{ij}=a_{ij}b_{ij}$, 则 c_{11} 在 $C=A\circ B$ 内的代数余子式为

$$C_{11}=d(C_{23\cdots n}^{23\cdots n})=\begin{vmatrix}c_{22}&c_{23}&\cdots&c_{2n}\\c_{32}&c_{33}&\cdots&c_{3n}\\\vdots&\vdots&&\vdots\\c_{n2}&c_{n3}&\cdots&c_{nn}\end{vmatrix},$$

故由归纳的假设有 $C_{11}\geq b_{22}\cdots b_{nn}A_{11}$，其中 $A_{11}=d(A_{23\cdots n}^{23\cdots n})$.

再令

$$A_0=\begin{pmatrix}a_{11}-|A|/A_{11}&a_{12}&\cdots&a_{1n}\\a_{21}&a_{22}&\cdots&a_{2n}\\\vdots&\vdots&&\vdots\\a_{n1}&a_{n2}&\cdots&a_{nn}\end{pmatrix},$$

则 $|A_0|=0$, $\det(A_{0\,r,r+1,\cdots,n}^{\,r,r+1,\cdots,n})>0(r\geq 2)$，故 A_0 为半正定的厄米特矩阵，因而 $A_0\circ B$ 是半正定的，随之有 $|A_0\circ B|=0$，即

$|A\circ B|-\dfrac{|A|}{A_{11}}\cdot b_{11}C_{11}=0$, 故

$$|A\circ B|=\frac{|A|}{A_{11}}\cdot b_{11}C_{11}\geqslant\frac{|A|}{A_{11}}\cdot b_{11}b_{22}\cdots b_{nn}A_{11}$$
$$=b_{11}b_{22}\cdots b_{nn}|A|,$$

定理 10 证完.

推论 1 A 与 B 都为正定或有一为半正定时，则 $|A\circ B|\geqslant|A||B|$.

证明 当 A、B 中有一是半正定时，上不等式之左右两端都是零，显然成立. 故只需考虑 A、B 都是正定的情形.

由定理 10, $|A\circ B|\geqslant b_{11}b_{22}\cdots b_{nn}\cdot|A|$; 但由 B 之正定性，利用定理 9(取 $r=1$)，就得到

$$|B|\leqslant b_{11}\cdot\begin{vmatrix}b_{22}&\cdots&b_{2n}\\\vdots&&\vdots\\b_{n2}&\cdots&b_{nn}\end{vmatrix},$$

再反复利用定理 9(总是取 $r=1$)，就可知

$$|B|\leqslant b_{11}\begin{vmatrix}b_{22}&\cdots&b_{2n}\\\vdots&&\vdots\\b_{n2}&\cdots&b_{nn}\end{vmatrix}\leqslant b_{11}b_{22}\begin{vmatrix}b_{33}&\cdots&b_{3n}\\\vdots&&\vdots\\b_{n3}&\cdots&b_{nn}\end{vmatrix}$$
$$\leqslant\cdots\leqslant b_{11}b_{22}\cdots b_{n-1,n-1}b_{nn},$$

于是，$|A\circ B|\geqslant b_{11}b_{22}\cdots b_{nn}\cdot|A|\geqslant|B|\cdot|A|$, 证完.

在证推论 1 的过程中，还证明了：

推论 2 若 $A=(a_{ik})$ 为正定厄米特矩阵，则必有 $|A|\leqslant a_{11}a_{22}\cdots a_{nn}$.

定理 11 设 $\varphi(x_1,\cdots,x_n)=\overline{X}'AX$ 是正定厄米特齐式，且对任 $X=\begin{pmatrix}x_1\\\vdots\\x_n\end{pmatrix}$ 恒有 $\varphi(x_1,\cdots,x_n)\leqslant\sum\limits_{i=1}^{n}|x_i|^2(=\overline{X}'X)$, 则必 $|A|\leqslant1$. 一般，若 $\psi(x_1,\cdots,x_n)=\overline{X}'BX$ 也是正定厄米特

齐式，且恒有 $\varphi(x_1, \cdots, x_n) \leqslant \psi(x_1, \cdots, x_n)$，则必 $|A| \leqslant |B|$.

证明 设 A 之 n 个特征值为 $\lambda_1, \lambda_2, \cdots, \lambda_n$，则 $|A| = \lambda_1\lambda_2\cdots\lambda_n$. 但对任一特征值 λ，取相应的任一特征向量 X 后，就有

$$\varphi(x_1, \cdots, x_n) = \overline{X}'AX = \lambda\overline{X}'X,$$

故再据题设 $\varphi(x_1, \cdots, x_n)$ 就有 $\lambda\overline{X}'X \leqslant \overline{X}'X$，因而就必有 $\lambda \leqslant 1$，于是 $|A| = \lambda_1\lambda_2\cdots\lambda_n \leqslant 1$.

一般，据引理 4 知有满秩变换 $X = PY$，使 $|P| = 1$ 且

$$\varphi(x_1, \cdots, x_n) = \overline{X}'AX = \overline{Y}'\overline{P}'APY = \sum_1^n \lambda_i \bar{y}_i y_i,$$

及
$$\psi(x_1, \cdots, x_n) = \overline{X}'BX = \overline{Y}'\overline{P}'BPY = \sum_1^n \mu_i \bar{y}_i y_i,$$

即
$$\overline{P}'AP = \operatorname{diag}\{\lambda_1, \lambda_2, \cdots, \lambda_n\},$$
$$\overline{P}'BP = \operatorname{diag}\{\mu_1, \mu_2, \cdots, \mu_n\},$$

λ_i 与 μ_i 都是正数. 故特取

$$y_1 = \cdots = y_{i-1} = y_{i+1} = \cdots = y_n = 0$$

而取 $y_i = 1$ 之 Y，这样所得的

$$X = PY \neq 0,$$

因之对这样的 X，据题设应有

$$\varphi(x_1, \cdots, x_n) = \lambda_i \leqslant \psi(x_1, \cdots, x_n) = \mu_i,$$

这说明了必是 $\lambda_i \leqslant \mu_i$，故

$$|A| = \lambda_1\lambda_2\cdots\lambda_n \leqslant \mu_1\mu_2\cdots\mu_n = |B|.$$

定理 11 证完.

关于酉矩阵，也有一些结果是很有意义的.

首先，若 $A = (a_{ik})$ 为酉矩阵，则 $\det A$ 之模为 1，A 当然是满秩的. 但酉矩阵 A 又具下述性质:

$$\begin{cases} a_{11}\cdot k\bar{a}_{i1}+a_{12}\cdot k\bar{a}_{i2}+\cdots+a_{1n}\cdot k\bar{a}_{in}=0, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{i-1,1}\cdot k\bar{a}_{i1}+a_{i-1,2}\cdot k\bar{a}_{i2}+\cdots+a_{i-1,n}\cdot k\bar{a}_{in}=0, \\ a_{i1}\cdot k\bar{a}_{i1}+a_{i2}\cdot k\bar{a}_{i2}+\cdots+a_{in}\cdot k\bar{a}_{in}=k, \\ a_{i+1,1}\cdot k\bar{a}_{i1}+a_{i+1,2}\cdot k\bar{a}_{i2}+\cdots+a_{i+1,n}\cdot k\bar{a}_{in}=0, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{n1}\cdot k\bar{a}_{i1}+a_{n2}\cdot k\bar{a}_{i2}+\cdots+a_{nn}\cdot k\bar{a}_{in}=0, \end{cases}$$

即

$$\boldsymbol{A}\cdot\begin{pmatrix} k\bar{a}_{i1} \\ k\bar{a}_{i2} \\ \vdots \\ k\bar{a}_{in} \end{pmatrix}=\begin{pmatrix} 0 \\ \vdots \\ 0 \\ k \\ 0 \\ \vdots \\ 0 \end{pmatrix}\to(\text{第 } i \text{ 行})$$

$$=\boldsymbol{B}_i,$$

式中 $k=\det\boldsymbol{A}$，这说明了方程 $\boldsymbol{AX}=\boldsymbol{B}_i$ 有解 $\begin{pmatrix} k\bar{a}_{i1} \\ \vdots \\ k\bar{a}_{in} \end{pmatrix}$。但显然 $\boldsymbol{AX}=\boldsymbol{B}_i$ 有解 $\begin{pmatrix} A_{i1} \\ A_{i2} \\ \vdots \\ A_{in} \end{pmatrix}$，故由于 $\det\boldsymbol{A}\neq 0$ 而据克莱姆定理之解的唯一性，就有

$$k\bar{a}_{ij}=A_{ij}\quad(i,\ j=1,\ 2,\ \cdots,\ n).$$

特别地，当 $\boldsymbol{A}$ 为实直交时，$k=\pm 1$，$\bar{a}_{ij}=a_{ij}$。于是证明了：

定理 12 酉矩阵 $\boldsymbol{A}=(a_{ij})$ 中每元 a_{ij} 的代数余子式 A_{ij} 总是等于行列式 $|\boldsymbol{A}|$ 乘以 a_{ij} 之共轭 $\bar{a}_{ij}$。特别，实直交矩阵 $\boldsymbol{A}$ 之每元之代数余子式或恒等于该元或恒等于该元之反号，全由 $\boldsymbol{A}$ 为正常或非正常的来决定。

一般，若不仅限于酉矩阵（n 级）$\boldsymbol{A}$ 之每元 a_{ij}，而考虑

$A=(a_{ij})$之任一个余子式(r 级)，又怎样呢？根据第2节，以 A之一切 r 级子式为元素作 A 之第 r 次的复合矩阵 $A^{(r)}$，则 $A^{(r)}$ 是 $\binom{n}{r}$ 级的矩阵，如令 $t=\binom{n}{r}$，则

$$A^{(r)}=\begin{pmatrix} c_{11} & c_{12} & \cdots & c_{1t} \\ c_{21} & c_{22} & \cdots & c_{2t} \\ \vdots & \vdots & & \vdots \\ c_{t1} & c_{t2} & \cdots & c_{tt} \end{pmatrix},$$

每 c_{ij} 是 A 的某个 r 级的子式．由复合矩阵之定义可知 $A^{(r)}$ 之每行中 t 个元素都是从 A 的某 r 个行中所取的 $t=\binom{n}{r}$个 r 级子式，例如令 c_{i1}，c_{i2}，…，c_{it} 是从 A 的第 i_1、i_2、…、i_r $(i_1<i_2<\cdots<i_r)$ 行所作的 t 个 r 级子式，而令 c_{ij} 在 A 中的代数余子式为 M_{ij}(因而是 A 中 $n-r$ 级子式乘以 (-1) 的某幂)．由拉普拉斯定理(行列式展开)得

$$\begin{cases} c_{11}M_{i1}+c_{12}M_{i2}+\cdots+c_{1t}M_{it}=0, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ c_{i1}M_{i1}+c_{i2}M_{i2}+\cdots+c_{it}M_{it}=k, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ c_{t1}M_{i1}+c_{t2}M_{i2}+\cdots+c_{tt}M_{it}=0, \end{cases}$$

但 $k=\det A^{(r)}=|A|^{\binom{n-1}{r-1}}$．另方面，$A$ 为酉矩阵也保证了 $A^{(r)}$ 是酉矩阵(第二节的定理9)，故

$$\begin{cases} c_{11}\cdot k\bar{c}_{i1}+c_{12}\cdot k\bar{c}_{i2}+\cdots+c_{1t}\cdot k\bar{c}_{it}=0, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ c_{i-1,1}\cdot k\bar{c}_{i1}+c_{i-1,2}\cdot k\bar{c}_{i2}+\cdots+c_{i-1,t}\cdot k\bar{c}_{it}=0, \\ c_{i1}\cdot k\bar{c}_{i1}+c_{i2}\cdot k\bar{c}_{i2}+\cdots+c_{it}\cdot k\bar{c}_{it}=k, \\ c_{i+1,1}\cdot k\bar{c}_{i1}+c_{i+1,2}\cdot k\bar{c}_{i2}+\cdots+c_{i+1,t}\cdot k\bar{c}_{it}=0, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ c_{t1}\cdot k\bar{c}_{i1}+c_{t2}\cdot k\bar{c}_{i2}+\cdots+c_{tt}\cdot k\bar{c}_{it}=0. \end{cases}$$

说明了
$$A^{(r)}\cdot X=B_i=\begin{pmatrix}0\\ \vdots\\ 0\\ k\\ 0\\ \vdots\\ 0\end{pmatrix}\to(\text{第 } i \text{ 行})$$

有解
$$\begin{pmatrix}k\bar{c}_{i1}\\ k\bar{c}_{i2}\\ \vdots\\ k\bar{c}_{it}\end{pmatrix} \text{ 与 } \begin{pmatrix}M_{i1}\\ M_{i2}\\ \vdots\\ M_{it}\end{pmatrix},$$

故从 $\det A^{(r)}\neq 0$ 而据克莱姆定理就有

$$k\bar{c}_{ij}=M_{ij}\quad (i,\ j=1,\ 2,\ \cdots,\ t).$$

特别地，当 A 为实直交时，$A^{(r)}$ 亦实直交，故 $\det A^{(r)}=\pm 1$，且 $\bar{c}_{ij}=c_{ij}$. 于是又证得了：

定理 13 n 级酉矩阵 $A=(a_{ij})$ 中每 r 级子式 M 的代数余子式 N 等于 $|A|^{\binom{n-1}{r-1}}$ 乘以 M 之共轭数 $\overline{M}$. 特别，实直交矩阵 A 之每子式或恒与其代数余子式相等或恒与其代数余子式反号.

第 5 节　全矩阵环的基

我们知道：设 C_{ij} 表示第 i 行第 j 列相交处的元素为 1，其余元素都为 0 的 n 级矩阵，那末 n^2 个矩阵 $C_{ij}(i,\ j=1,\ 2,\ \cdots,\ n)$ 具有性质：

$$C_{ij}C_{kl}=\begin{cases}C_{il}, & \text{当 } j=k \text{ 时};\\ O, & \text{当 } j\neq k \text{ 时}.\end{cases}$$

现在提出这样一个问题：如果有 n^2 个 n 级矩阵 $\boldsymbol{F}_{ij}(i, j=1, 2, \cdots, n)$ 也具有类似于 $\boldsymbol{C}_{ij}$ 间所具有之性质：

$$\boldsymbol{F}_{ij}\boldsymbol{F}_{kl}=\begin{cases}\boldsymbol{F}_{il}, & \text{当 } j=k \text{ 时;}\\ \boldsymbol{O}, & \text{当 } j\neq k \text{ 时,}\end{cases}$$

那末 n^2 个 $\boldsymbol{F}_{ij}$ 和 n^2 个 $\boldsymbol{C}_{ij}$ 的关系怎样呢†？

首先，若有某 $\boldsymbol{F}_{st}=\boldsymbol{O}$，则对任 i、j 应有

$$\boldsymbol{F}_{ij}=\boldsymbol{F}_{is}\boldsymbol{F}_{st}\boldsymbol{F}_{tj}=\boldsymbol{O},$$

这说明了 n^2 个 $\boldsymbol{F}_{ij}$ 中若有一个为零矩阵，则 n^2 个 $\boldsymbol{F}_{ij}$ 都是零矩阵. 因而，今后可假定每 $\boldsymbol{F}_{ij}\neq\boldsymbol{O}$.

今考虑 n 维向量空间 $\mathfrak{R}_n$，取其一组基底 $\boldsymbol{e}_1, \boldsymbol{e}_2, \cdots, \boldsymbol{e}_n$. 并令 $\mathscr{T}_{ij}$ 表示对应于这组基底之矩阵为 $\boldsymbol{F}_{ij}$ 的 $\mathfrak{R}_n$ 之线性变换，即

$$\begin{pmatrix}\boldsymbol{e}_1\\ \boldsymbol{e}_2\\ \vdots\\ \boldsymbol{e}_n\end{pmatrix}\mathscr{T}_{ij}=\boldsymbol{F}_{ij}\begin{pmatrix}\boldsymbol{e}_1\\ \boldsymbol{e}_2\\ \vdots\\ \boldsymbol{e}_n\end{pmatrix}, \quad (i, j=1, 2, \cdots, n).$$

从 $\boldsymbol{F}_{11}\neq\boldsymbol{O}$ 及 $\boldsymbol{F}_{11}^2=\boldsymbol{F}_{11}$ 知 $\boldsymbol{F}_{11}$ 为幂等矩阵，故特征值有 1 及 0，$\lambda(\lambda-1)$ 为零化多项式又说明了 $\boldsymbol{F}_{11}$ 之最小多项式无重根，因而 $\boldsymbol{F}_{11}\sim\operatorname{diag}\{1, \cdots, 1, 0, \cdots, 0\}$，且至少有一个 1 出现(否则将导致 $\boldsymbol{F}_{11}=\boldsymbol{O}$，显非所许)，同时也至少有一个 0 出现(否则有 $\boldsymbol{F}_{11}=\boldsymbol{E}$，随而 $\boldsymbol{O}=\boldsymbol{F}_{22}\boldsymbol{F}_{11}=\boldsymbol{F}_{22}\boldsymbol{E}=\boldsymbol{F}_{22}$，也不允许). 这样，可取 $\mathfrak{R}_n$ 的某适当基底，使 $\mathscr{T}_{11}$ 对这组基底之矩阵即为这对角阵 $\operatorname{diag}\{1, \cdots, 1, 0, \cdots, 0\}$，故若令此基底的第一个向量为 $\boldsymbol{f}_1$，则当然有 $\boldsymbol{f}_1\neq\boldsymbol{0}$ 且 $\boldsymbol{f}_1\mathscr{T}_{11}=\boldsymbol{f}_1$.

同理，考虑 $\boldsymbol{F}_{22}, \cdots, \boldsymbol{F}_{nn}$ 就会分别有向量 $\boldsymbol{f}_2, \cdots, \boldsymbol{f}_n$ 使

† 这个问题由我校代数研究生班的同志们共同讨论解决的，他们解决得很好，今摘录于下.

$\boldsymbol{f}_i\mathscr{T}_{ii}=\boldsymbol{f}_i$ 且 $\boldsymbol{f}_i\neq\boldsymbol{0}(i=2,\cdots,n)$.

这就找出了 $\mathfrak{R}_n$ 的 n 个非零的向量 $\boldsymbol{f}_1,\boldsymbol{f}_2,\cdots,\boldsymbol{f}_n$. 下面来证明 $\boldsymbol{f}_1,\boldsymbol{f}_2,\cdots,\boldsymbol{f}_n$ 是线性无关的: 因若 $\boldsymbol{f}_1,\boldsymbol{f}_2,\cdots,\boldsymbol{f}_n$ 线性相关,则有一(例如 $\boldsymbol{f}_1$)可写为 $\boldsymbol{f}_2,\cdots,\boldsymbol{f}_n$ 之线性组合,如

$$\boldsymbol{f}_1=\sum_{j=2}^{n}k_j\boldsymbol{f}_j,$$

于是

$$\boldsymbol{f}_1=\boldsymbol{f}_1\mathscr{T}_{11}=\sum_{j=2}^{n}k_j(\boldsymbol{f}_j\mathscr{T}_{11}); \tag{8}$$

但注意到 $j\neq i$ 时, $\boldsymbol{F}_{jj}\boldsymbol{F}_{ii}=\boldsymbol{O}$, 故 $\mathscr{T}_{jj}\mathscr{T}_{ii}$ 使 $\mathfrak{R}_n$ 之每向量都变为零向量,因而有 $\boldsymbol{f}_j\mathscr{T}_{ii}=(\boldsymbol{f}_j\mathscr{T}_{jj})\mathscr{T}_{ii}=\boldsymbol{f}_j(\mathscr{T}_{jj}\mathscr{T}_{ii})=\boldsymbol{0}$, 因而再据(8)式得 $\boldsymbol{f}_1=0$, 这就和 $\boldsymbol{f}_1\neq\boldsymbol{0}$ 矛盾.

$\boldsymbol{f}_1,\boldsymbol{f}_2,\cdots,\boldsymbol{f}_n$ 既线性无关, 说明了 $\boldsymbol{f}_1,\boldsymbol{f}_2,\cdots,\boldsymbol{f}_n$ 可为 $\mathfrak{R}_n$ 之一组基底. 且由上面一段之证明,得知

$$\boldsymbol{f}_i\mathscr{T}_{ii}=\boldsymbol{f}_i,\quad \boldsymbol{f}_j\mathscr{T}_{ii}=\boldsymbol{0}\quad(\text{当 } j\neq i \text{ 时}), \tag{9}$$

这表明对于基底 $\boldsymbol{f}_1,\boldsymbol{f}_2,\cdots,\boldsymbol{f}_n$, 线性变换 $\mathscr{T}_{ii}$ 所对应之矩阵为 $\boldsymbol{C}_{ii}$. 显然, 不论 $k_1,k_2,\cdots,k_n$ 是任何 n 个非零的数, 则 $k_1\boldsymbol{f}_1,k_2\boldsymbol{f}_2,\cdots,k_n\boldsymbol{f}_n$ 亦为 $\mathfrak{R}_n$ 之一组基底, 且 $(k_i\boldsymbol{f}_i)\mathscr{T}_{ii}=(k_i\boldsymbol{f}_i)$ 及 $(k_j\boldsymbol{f}_j)\mathscr{T}_{ii}=\boldsymbol{0}$(当 $j\neq i$ 时) 又说明了线性变换 $\mathscr{T}_{ii}$ 所对应之矩阵(关于基底 $k_1\boldsymbol{f}_1,k_2\boldsymbol{f}_2,\cdots,k_n\boldsymbol{f}_n$) 仍为 $\boldsymbol{C}_{ii}$.

再来讨论线性变换 $\mathscr{T}_{ij}$(当 $i\neq j$ 时)对基底 $\boldsymbol{f}_1,\boldsymbol{f}_2,\cdots,\boldsymbol{f}_n$ 的作用: 显然,当 $k\neq i$ 时有

$$\boldsymbol{f}_k\mathscr{T}_{ij}=\boldsymbol{f}_k(\mathscr{T}_{ii}\mathscr{T}_{ij})=(\boldsymbol{f}_k\mathscr{T}_{ii})\mathscr{T}_{ij}=\boldsymbol{0}\mathscr{T}_{ij}=\boldsymbol{0}, \tag{10}$$

但若令 $\boldsymbol{f}_i\mathscr{T}_{ij}=\lambda_{i1}^{(j)}\boldsymbol{f}_1+\lambda_{i2}^{(j)}\boldsymbol{f}_2+\cdots+\lambda_{in}^{(j)}\boldsymbol{f}_n$, 就有

$$\begin{aligned}\boldsymbol{f}_i\mathscr{T}_{ij}&=\boldsymbol{f}_i(\mathscr{T}_{ij}\mathscr{T}_{jj})=(\boldsymbol{f}_i\mathscr{T}_{ij})\mathscr{T}_{jj}=\sum_{k=1}^{n}(\lambda_{ik}^{(j)}\boldsymbol{f}_k)\mathscr{T}_{jj}\\&=\sum_{k=1}^{n}\lambda_{ik}^{(j)}(\boldsymbol{f}_k\mathscr{T}_{jj})=\lambda_{ij}^{(j)}\boldsymbol{f}_j,\end{aligned} \tag{11}$$

且必 $\lambda_{ij}^{(j)}\neq 0$(否则,就有 $\mathscr{T}_{ij}$ 使 $\boldsymbol{f}_1,\boldsymbol{f}_2,\cdots,\boldsymbol{f}_n$ 都变为零,从而

$\mathscr{T}_{ij}=\mathscr{O}$, 显非所许).

合并(9), (10), (11)三个式子,可概括写成

$$\left.\begin{aligned}&\boldsymbol{f}_i\mathscr{T}_{ij}=m_{ij}\boldsymbol{f}_j, \quad m_{ij}\neq 0;\\&\boldsymbol{f}_k\mathscr{T}_{ij}=\mathbf{0} \quad (\text{当 } k\neq i \text{ 时}),\end{aligned}\right\}\tag{12}$$

式中 i、$j=1, 2, \cdots, n$, 而 i 与 j 相等与否是无关的, 且有 $m_{ii}=1$. 于是有

$$\begin{aligned}m_{ik}\boldsymbol{f}_k=\boldsymbol{f}_i\mathscr{T}_{ik}&=\boldsymbol{f}_i(\mathscr{T}_{ij}\mathscr{T}_{jk})=(\boldsymbol{f}_i\mathscr{T}_{ij})\mathscr{T}_{jk}\\&=m_{ij}\boldsymbol{f}_j\mathscr{T}_{jk}=m_{ij}m_{jk}\boldsymbol{f}_k,\end{aligned}$$

不得不有

且特有

$$\left.\begin{aligned}&m_{ik}=m_{ij}m_{jk},\\&m_{ij}m_{ji}=m_{ii}=1.\end{aligned}\right\}\tag{13}$$

因而矩阵

$$\boldsymbol{M}=\begin{pmatrix}m_{11} & m_{12} & \cdots & m_{1n}\\ m_{21} & m_{22} & \cdots & m_{2n}\\ \vdots & \vdots & & \vdots\\ m_{n1} & m_{n2} & \cdots & m_{nn}\end{pmatrix}$$

中任一个二级子式(利用(13)式)为

$$\begin{aligned}\Delta=\begin{vmatrix}m_{ik} & m_{il}\\ m_{jk} & m_{jl}\end{vmatrix}&=m_{ik}m_{jl}-m_{il}m_{jk}=m_{ik}m_{jl}-m_{ik}m_{kl}m_{jk}\\&=m_{ik}m_{jl}-m_{ik}m_{jk}m_{kl}=m_{ik}m_{jl}-m_{ik}m_{jl}=0,\end{aligned}$$

说明了 $\boldsymbol{M}$ 的秩等于 1, 故必有二组全非零的 n 个数 $a_1, a_2, \cdots, a_n$ 及 $b_1, b_2, \cdots, b_n$ 使 $m_{ij}=a_ib_j$, 于是 $1=m_{ii}=a_ib_i$ 表示着 $b_i=a_i^{-1}$.

再令 $\boldsymbol{g}_i=b_i\boldsymbol{f}_i$, 故将 b_i 看作上述的 k_i, 则 n 个向量 $\boldsymbol{g}_1, \boldsymbol{g}_2, \cdots, \boldsymbol{g}_n$ 构成 $\mathfrak{R}_n$ 之一基底; 且关于这组基底, 线性变换 $\mathscr{T}_{ii}$ 所对应之矩阵为 $\boldsymbol{C}_{ii}$(上面已谈过); 但在 $k\neq i$ 时有(利用(12)式):

$$\boldsymbol{g}_k\mathscr{T}_{ij}=b_k\boldsymbol{f}_k\mathscr{T}_{ij}=\boldsymbol{0},$$

而 $$\boldsymbol{g}_i\mathscr{T}_{ij}=b_i\boldsymbol{f}_i\mathscr{T}_{ij}=b_im_{ij}\boldsymbol{f}_j=b_ia_ib_j\boldsymbol{f}_j=\boldsymbol{g}_j,$$

又说明了 $\mathscr{T}_{ij}$ 所对应的矩阵是 $\boldsymbol{C}_{ij}$(当 $j\neq i$ 时).

这说明了：关于 $\mathfrak{R}_n$ 之基底 $\boldsymbol{g}_1, \boldsymbol{g}_2, \cdots, \boldsymbol{g}_n$, 线性变换 $\mathscr{T}_{ij}$ 对应之矩阵为 $\boldsymbol{C}_{ij}(i, j=1, 2, \cdots, n)$. 若从 $\boldsymbol{e}_1, \cdots, \boldsymbol{e}_n$ 到 $\boldsymbol{g}_1, \boldsymbol{g}_2, \cdots, \boldsymbol{g}_n$ 之变换矩阵为 $\boldsymbol{P}$, 即

$$\boldsymbol{g}_i=\sum_{j=1}^{n}p_{ij}\boldsymbol{e}_j,$$

则 $\boldsymbol{P}=(p_{ij})$ 是满秩的，且有

$$\begin{pmatrix}\boldsymbol{g}_1\\ \boldsymbol{g}_2\\ \vdots\\ \boldsymbol{g}_n\end{pmatrix}\mathscr{T}_{ij}=\boldsymbol{C}_{ij}\begin{pmatrix}\boldsymbol{g}_1\\ \boldsymbol{g}_2\\ \vdots\\ \boldsymbol{g}_n\end{pmatrix};$$

但 $$\begin{pmatrix}\boldsymbol{g}_1\\ \vdots\\ \boldsymbol{g}_n\end{pmatrix}\mathscr{T}_{ij}=\boldsymbol{P}\begin{pmatrix}\boldsymbol{e}_1\\ \vdots\\ \boldsymbol{e}_n\end{pmatrix}\mathscr{T}_{ij}=\boldsymbol{P}\boldsymbol{F}_{ij}\begin{pmatrix}\boldsymbol{e}_1\\ \vdots\\ \boldsymbol{e}_n\end{pmatrix},$$

$$\boldsymbol{C}_{ij}\begin{pmatrix}\boldsymbol{g}_1\\ \vdots\\ \boldsymbol{g}_n\end{pmatrix}=\boldsymbol{C}_{ij}\boldsymbol{P}\begin{pmatrix}\boldsymbol{e}_1\\ \vdots\\ \boldsymbol{e}_n\end{pmatrix},$$

故必 $$\boldsymbol{P}\boldsymbol{F}_{ij}=\boldsymbol{C}_{ij}\boldsymbol{P},$$

即 $$\boldsymbol{F}_{ij}=\boldsymbol{P}^{-1}\boldsymbol{C}_{ij}\boldsymbol{P}\quad(i, j=1, 2, \cdots, n).$$

如果又有满秩矩阵 $\boldsymbol{Q}$ 使

$$\boldsymbol{F}_{ij}=\boldsymbol{Q}^{-1}\boldsymbol{C}_{ij}\boldsymbol{Q},$$

则 $$\boldsymbol{P}^{-1}\boldsymbol{C}_{ij}\boldsymbol{P}=\boldsymbol{Q}^{-1}\boldsymbol{C}_{ij}\boldsymbol{Q},\quad \boldsymbol{Q}\boldsymbol{P}^{-1}\boldsymbol{C}_{ij}=\boldsymbol{C}_{ij}\boldsymbol{Q}\boldsymbol{P}^{-1},$$

即 $\boldsymbol{Q}\boldsymbol{P}^{-1}$ 与 $\boldsymbol{C}_{ij}$ 因之与任何矩阵相乘可交换，故 $\boldsymbol{Q}\boldsymbol{P}^{-1}$ 为纯量矩阵 $k\boldsymbol{E}$, 因而 $\boldsymbol{Q}=k\boldsymbol{P}$.

这就证得了：

定理 设有 n^2 个 n 级矩阵 $\boldsymbol{F}_{ij}(i, j=1, 2\cdots, n)$, 那末 n^2

个 F_{ij} 满足关系式

$$F_{ij}F_{kl}=\begin{cases} F_{il}, & \text{当 } j=k \text{ 时;} \\ O, & \text{当 } j\neq k \text{ 时} \end{cases}$$

的充要条件是有一满秩矩阵 P, 使得

$$F_{ij}=P^{-1}C_{ij}P.$$

并且 P 除相差一常数因子外还唯一地存在.

下面介绍上述定理的第二种证法. 首先, 若 n^2个$F_{ij}(i, j=1, 2, \cdots, n)$ 中有一个为零, 那末这 n^2 个 F_{ij} 必全都是零. 这与上证法一样. 故今后可假定每 $F_{ij}\neq O$.

其次, 仅先考虑 n 个 $F_{ii}(i=1, 2, \cdots, n)$. $F_{ii}^2=F_{ii}$ 说明了 $F_{ii}(F_{ii}-E)=O$. 可以证明 $F_{ii}-E\neq O$, 今用反证法证之: 若 $F_{ii}=E$, 则 $j\neq i$ 时有 $O=F_{jj}F_{ii}=F_{jj}E=F_{jj}$, 因而 $F_{ij}=F_{ij}F_{jj}=O$, $F_{ji}=F_{jj}F_{ji}=O$, $E=F_{ii}=F_{ij}F_{ji}=O$, 此即矛盾(或直接得到 $O=F_{jj}$ 后说明矛盾亦可). 这说明了任何 $F_{ii}\neq O$ 及 $F_{ii}-E\neq O$, 故由其积为零则知 $F_{ii}-E$ 为降秩的, 于是$(F_{ii}-E)X=\mathbf{0}$ 有非零解, 取其任一非零解 X_i, 就得到了 n 个 n 元列向量 $X_i(i=1, \cdots, n)$, 它们具有下列性质:

$$F_{ii}X_j=\begin{cases} X_i, & \text{当 } j=i \text{ 时;} \\ \mathbf{0}, & \text{当 } j\neq i \text{ 时.} \end{cases} \tag{14}$$

事实上, 当 $j=i$ 时的结论是由 X_i 之取法而得的; 当 $j\neq i$ 时有 $F_{ii}X_j=F_{ii}(F_{jj}X_j)=(F_{ii}F_{jj})X_j=OX_j=\mathbf{0}$.

今可证明 $X_1, X_2, \cdots, X_n$ 是线性无关的. 因从

$$\sum_{i=1}^{n}a_iX_i=\mathbf{0}$$

之左右两端均用 F_{kk} 去左乘面利用(14)时, 就得

$$a_kX_k=\mathbf{0},$$

因 $X_k\neq\mathbf{0}$ 从面有 $a_k=0(k=1, 2, \cdots, n)$, 即 $X_1, X_2, \cdots, X_n$

线性无关. 由是得知任一个 n 元列向量都可写为 $\boldsymbol{X}_1, \boldsymbol{X}_2, \cdots, \boldsymbol{X}_n$ 之线性组合.

据此, 可以证明 $\boldsymbol{F}_{ii}\boldsymbol{X}=\boldsymbol{X}$ 之解空间是一维的, 即 $r_{(\boldsymbol{F}_{ii}-\boldsymbol{E})}=n-1$: 事实上, 若

$$\boldsymbol{Y}=\sum_{k=1}^{n}\lambda_k\boldsymbol{X}_k$$

是它的一解, 则利用(14)得

$$\sum_{k=1}^{n}\lambda_k\boldsymbol{X}_k=\boldsymbol{Y}=\boldsymbol{F}_{ii}\boldsymbol{Y}=\boldsymbol{F}_{ii}\cdot\sum_{k=1}^{n}\lambda_k\boldsymbol{X}_k=\lambda_i\boldsymbol{X}_i,$$

即

$$\sum_{\substack{k=1\\k\neq i}}^{n}\lambda_k\boldsymbol{X}_k=\boldsymbol{0};$$

故由 $\boldsymbol{X}_k\,(1\leqslant k\leqslant n,\ k\neq i)$ 线性无关即得 $\lambda_k=0\,(1\leqslant k\leqslant n,\ k\neq i)$, 可见 $\boldsymbol{F}_{ii}\boldsymbol{X}=\boldsymbol{X}$ 之解只能是 $\boldsymbol{X}=\lambda\boldsymbol{X}_i$ 形, 即解空间为一维的.

再作满秩矩阵 $\boldsymbol{P}=(\boldsymbol{Y}_1, \boldsymbol{Y}_2, \cdots, \boldsymbol{Y}_n)$ 之 n 个列向量 $\boldsymbol{Y}_i$ 使具有下述性质

$$\boldsymbol{F}_{ij}\boldsymbol{Y}_k=\begin{cases}\boldsymbol{Y}_i, & \text{当 } j=k \text{ 时};\\ \boldsymbol{0}, & \text{当 } j\neq k \text{ 时}.\end{cases}\tag{15}$$

具体作法于下:

先令 $\boldsymbol{Y}_1=\boldsymbol{X}_1$, 然后由 $\boldsymbol{F}_{11}(\boldsymbol{F}_{1j}\boldsymbol{X}_j)=(\boldsymbol{F}_{11}\boldsymbol{F}_{1j})\boldsymbol{X}_j=\boldsymbol{F}_{1j}\boldsymbol{X}_j$ $(j=2, \cdots, n)$而据上而一段所云得知 $\boldsymbol{F}_{1j}\boldsymbol{X}_j=c_j\boldsymbol{Y}_1$, 并可断言 $c_j\neq0$ (因不然就有 $\boldsymbol{F}_{1j}\boldsymbol{X}_j=\boldsymbol{0}\Rightarrow\boldsymbol{X}_j=\boldsymbol{F}_{jj}\boldsymbol{X}_j=\boldsymbol{F}_{j1}\boldsymbol{F}_{1j}\boldsymbol{X}_j=\boldsymbol{0}$, 矛盾). 因而, 可令 $\boldsymbol{Y}_j=\dfrac{1}{c_j}\boldsymbol{X}_j\,(j=1, 2, \cdots, n;$ 但规定 $c_1=1)$. 由 $\boldsymbol{X}_1, \boldsymbol{X}_2, \cdots, \boldsymbol{X}_n$ 线性无关即知 $\boldsymbol{Y}_1, \boldsymbol{Y}_2, \cdots, \boldsymbol{Y}_n$ 线性无关, 故有满秩矩阵 $\boldsymbol{P}=(\boldsymbol{Y}_1, \boldsymbol{Y}_2, \cdots, \boldsymbol{Y}_n)$. 且据 $\boldsymbol{Y}_i$ 之作法易知

$$\boldsymbol{F}_{1j}\boldsymbol{Y}_j=\boldsymbol{Y}_1\quad(j=2, \cdots, n),$$

且

$$\boldsymbol{F}_{ii}\boldsymbol{Y}_i=\boldsymbol{Y}_i\quad(i=1, 2, \cdots, n).$$

从而，可以验证：

$$F_{ij}Y_k=\begin{cases}F_{i1}F_{1j}Y_j=F_{i1}Y_1=F_{i1}F_{1i}Y_i\\ \qquad\qquad =F_{ii}Y_i=Y_i, & \text{当 } j=k \text{ 时；}\\ F_{ij}F_{kk}Y_k=OY_k=0, & \text{当 } j\neq k \text{ 时．}\end{cases}$$

即(15)成立．

由直接验证又知

$$PC_{ij}=(Y_1,\ Y_2,\ \cdots,\ Y_n)\cdot C_{ij}=(0,\ \cdots,\ 0,\ \underset{\substack{\downarrow\\ \text{第 } j \text{ 列}}}{Y_i},\ 0,\ \cdots,\ 0),$$

及

$$\begin{aligned}F_{ij}P&=F_{ij}\cdot(Y_1,\ Y_2,\ \cdots,\ Y_n)=(F_{ij}Y_1,\ F_{ij}Y_2,\ \cdots,\ F_{ij}Y_n)\\&=(0,\ \cdots,\ 0,\ \underset{\substack{\downarrow\\ \text{第 } j \text{ 列}}}{Y_i},\ 0,\ \cdots,\ 0)\end{aligned}$$

(后者要利用(15)式)．故

$$PC_{ij}=F_{ij}P,$$

即

$$C_{ij}=P^{-1}F_{ij}P.$$

即证得了定理中 P 之存在性．

下面给出第 485 页定理的第三种证法．n^2 个 F_{ij} 中如有一为零，则 n^2 个 F_{ij} 必都为零矩阵，证法一样．故今后可假定每个 $F_{ij}\neq O$．

先讨论 $F_{11}, F_{22}, \cdots, F_{nn}$ 这 n 个矩阵．因 $i\neq j$ 时 $F_{ii}F_{jj}=O=F_{jj}F_{ii}$，且每 F_{ii} 又与对角矩阵相似而对角线上的数不外乎是 1 与 0(因 $F_{ii}^2=F_{ii}$ 说明了 F_{ii} 之最小多项式无重根且特征值只能为 1 与 0)，故必有一满秩矩阵 P 使 $F_{11}, F_{22}, \cdots, F_{nn}$ 同时变为对角矩阵†，即

---

† 一般，有：若 $A_1, A_2, \cdots, A_k$ 是 k 个两两相乘为可交换的矩阵，即 $A_iA_j=A_jA_i$，且每 A_i 又与对角矩阵相似，那末必可用同一个满秩矩阵 P 使 $P^{-1}A_iP$ 为对角形$(i=1, \cdots, k)$．这准备在后面作为一引理给以证明．

$$P^{-1}F_{ii}P=\operatorname{diag}\{\delta_1^{(i)},\ \delta_2^{(i)},\ \cdots,\ \delta_n^{(i)}\}.$$

但每 $\delta_j^{(i)}=1$ 或 $0(i=1,\ 2,\ \cdots,\ n)$. 由于

$$\begin{aligned}O&=P^{-1}F_{ii}P\cdot P^{-1}F_{jj}P\\&=\operatorname{diag}\{\delta_1^{(i)}\delta_1^{(j)},\ \delta_2^{(i)}\delta_2^{(j)},\ \cdots,\ \delta_n^{(i)}\delta_n^{(j)}\},\end{aligned}$$

则知 $i\neq j$ 时对任一 $k=1,\ 2,\ \cdots,\ n$ 皆有 $\delta_k^{(i)}\delta_k^{(j)}=0$, 于是 $\delta_k^{(1)}$, $\delta_k^{(2)}$, $\cdots$, $\delta_k^{(n)}$ 中只能有一个且只一个等于 1(对任意的 $k=1$, $2,\cdots,n$); 为简单计且不损一般性, 令 $\delta_1^{(1)}=1,\delta_2^{(1)}=\cdots=\delta_n^{(1)}=0$; $\delta_1^{(2)}=\delta_3^{(2)}=\cdots=\delta_n^{(2)}=0,\delta_2^{(2)}=1$; $\cdots$; $\delta_1^{(n)}=\cdots=\delta_{n-1}^{(n)}=0,\delta_n^{(n)}=1$; 即 $P^{-1}F_{ii}P=C_{ii}(i=1,\ 2,\ \cdots,\ n)$.

再考虑 $P^{-1}F_{ij}P(i\neq j)$. 令 $P^{-1}F_{ij}P=(f_{ij})$, 易知

$$\begin{aligned}P^{-1}F_{ij}P&=P^{-1}F_{ii}P\cdot P^{-1}F_{ij}P\cdot P^{-1}F_{jj}P\\&=C_{ii}\cdot P^{-1}F_{ij}P\cdot C_{jj}\\&=\begin{pmatrix}O\\f_{i1}f_{i2}\cdots f_{in}\\O\end{pmatrix}\cdot C_{jj}=f_{ij}C_{ij},\end{aligned}$$

故由 $F_{ij}\neq O$ 知 $f_{ij}\neq 0$. 这说明对任何的 $i,\ j=1,\ 2,\ \cdots,\ n$ 恒有

$$P^{-1}F_{ij}P=\lambda_{ij}C_{ij},\tag{16}$$

$\lambda_{ij}\neq 0$ 且 $\lambda_{ii}=1$.

由于 $P^{-1}F_{ij}P\cdot P^{-1}F_{ji}P=P^{-1}F_{ii}P=C_{ii}$ 即 $\lambda_{ij}\lambda_{ji}C_{ii}=C_{ii}$, 又得 $\lambda_{ij}\lambda_{ji}=1=\lambda_{ii}$ 且一般有 $\lambda_{ij}\lambda_{jt}=\lambda_{it}$. 令 P 之 n 个列向量为 $X_1,\ X_2,\ \cdots,\ X_n$, 即 $P=(X_1,\ X_2,\ \cdots,\ X_n)$, 则从(16)式得

$$F_{ij}P=\lambda_{ij}PC_{ij}=\lambda_{ij}\cdot(0,\ \cdots,\ 0,\ \underset{(\text{第 }j\text{ 列})}{X_i},\ 0,\ \cdots,\ 0),$$

从而有

$$\begin{cases}F_{ij}X_k=0 & (\text{当 }k\neq j\text{ 时}),\\F_{ij}X_j=\lambda_{ij}X_i & (i,\ j=1,\ 2,\ \cdots,\ n).\end{cases}\tag{17}$$

于是由$\begin{pmatrix} \lambda_{11} & \lambda_{12} & \cdots & \lambda_{1n} \\ \lambda_{21} & \lambda_{22} & \cdots & \lambda_{2n} \\ \vdots & \vdots & & \vdots \\ \lambda_{n1} & \lambda_{n2} & \cdots & \lambda_{nn} \end{pmatrix}$中任二级子式

$$\begin{vmatrix} \lambda_{ik} & \lambda_{il} \\ \lambda_{jk} & \lambda_{jl} \end{vmatrix} = \lambda_{ik}\lambda_{jl} - \lambda_{il}\lambda_{jk} = \lambda_{ik}\lambda_{jl} - \lambda_{ik}\lambda_{kl}\lambda_{jl}\lambda_{lk}$$

$$= \lambda_{ik}\lambda_{jl} - \lambda_{ik}\lambda_{jl} = 0,$$

知使 $\lambda_{ij} = a_i b_j$ 成立的两组不为零的 n 个数 $a_1, a_2, \cdots, a_n$ 与 $b_1, b_2, \cdots, b_n$ 是存在的. 故令 $\boldsymbol{Y}_i = \frac{1}{b_i}\boldsymbol{X}_i (i=1, \cdots, n)$时, 则以 $\boldsymbol{Y}_1, \cdots, \boldsymbol{Y}_n$ 为 n 个列向量所作成的矩阵 $\boldsymbol{Q} = (\boldsymbol{Y}_1, \boldsymbol{Y}_2, \cdots, \boldsymbol{Y}_n)$是满秩的, 且因(利用(17)式)

$$\begin{cases} \boldsymbol{F}_{ij}\boldsymbol{Y}_k = \boldsymbol{F}_{ij}\boldsymbol{X}_k/b_k = \boldsymbol{0} \quad (k \neq j), \\ \boldsymbol{F}_{ij}\boldsymbol{Y}_j = \frac{1}{b_j}\boldsymbol{F}_{ij}\boldsymbol{X}_j = \frac{1}{b_j}\lambda_{ij}\boldsymbol{X}_i \\ \qquad = \frac{1}{b_j}a_i b_j \boldsymbol{X}_i = a_i\boldsymbol{X}_i = \frac{\boldsymbol{X}_i}{b_i} = \boldsymbol{Y}_i, \end{cases}$$

即 $\boldsymbol{F}_{ij}\boldsymbol{Q} = \boldsymbol{Q}\boldsymbol{C}_{ij}$, 或 $\boldsymbol{Q}^{-1}\boldsymbol{F}_{ij}\boldsymbol{Q} = \boldsymbol{C}_{ij}$, 这也证明了定理中 $\boldsymbol{P}$ 之存在性.

于是, 为完成第三种证法, 还需证明下引理.

引理 设 $\boldsymbol{A}_1, \boldsymbol{A}_2, \cdots, \boldsymbol{A}_k$ 是 k 个两两相乘为可交换的矩阵($\boldsymbol{A}_i\boldsymbol{A}_j = \boldsymbol{A}_j\boldsymbol{A}_i$), 且每 $\boldsymbol{A}_i$ 又与对角矩阵相似. 那末存在同一个满秩矩阵 $\boldsymbol{P}$, 使 $\boldsymbol{P}^{-1}\boldsymbol{A}_i\boldsymbol{P}$ 为对角矩阵($i=1, 2, \cdots, k$).

证明 先讨论 $k=2$ 的场合, 即 $\boldsymbol{AB} = \boldsymbol{BA}$, 且 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 都与对角矩阵相似. 因有满秩矩阵 $\boldsymbol{Q}_1$ 使

$$\boldsymbol{Q}_1^{-1}\boldsymbol{A}\boldsymbol{Q}_1 = \mathrm{diag}\{\lambda_1\boldsymbol{E}_1, \lambda_2\boldsymbol{E}_2, \cdots, \lambda_t\boldsymbol{E}_t\},$$

$\lambda_i \neq \lambda_j$ ($i \neq j$ 时), $\boldsymbol{E}_i$ 为 n_i 级单位矩阵 ($i=1, \cdots, t$), 因而

$\sum\limits_{i=1}^{t} n_i = n$ 为 $\boldsymbol{A}(\boldsymbol{B})$ 之级数，故从 $\boldsymbol{Q}_1^{-1}\boldsymbol{B}\boldsymbol{Q}_1$ 与 $\boldsymbol{Q}_1^{-1}\boldsymbol{A}\boldsymbol{Q}_1$ 可交换，得知

$$\boldsymbol{Q}_1^{-1}\boldsymbol{B}\boldsymbol{Q}_1=\begin{pmatrix} \boldsymbol{B}_1 & & & 0 \\ & \boldsymbol{B}_2 & & \\ & & \ddots & \\ 0 & & & \boldsymbol{B}_t \end{pmatrix}$$

为对角分块矩阵，$\boldsymbol{B}_i$ 是 n_i 级的．

由于 $\boldsymbol{B}$ 与对角矩阵相似，知 $\lambda\boldsymbol{B}-\boldsymbol{B}$ 的初等因子都是 λ 的一次式；但 $\lambda\boldsymbol{E}_i-\boldsymbol{B}_i$ 之初等因子 $(i=1, \cdots, t)$ 的全部即为 $\lambda\boldsymbol{E}-\boldsymbol{B}$ 之初等因子，故每 $\lambda\boldsymbol{B}_i-\boldsymbol{B}_i$ 的初等因子也是 λ 的一次式，因而每 $\boldsymbol{B}_i(i=1, \cdots, t)$ 也必与对角矩阵相似，即有 $\boldsymbol{R}_i$ $(i=1, 2, \cdots, t)$ 使 $\boldsymbol{R}_i^{-1}\boldsymbol{B}_i\boldsymbol{R}_i$ 为对角阵；于是，令

$$\boldsymbol{Q}_2=\begin{pmatrix} \boldsymbol{R}_1 & & & 0 \\ & \boldsymbol{B}_2 & & \\ & & \ddots & \\ 0 & & & \boldsymbol{B}_t \end{pmatrix}$$

时，则有

$$\boldsymbol{Q}_2^{-1}\boldsymbol{Q}_1^{-1}\boldsymbol{B}\boldsymbol{Q}_1\boldsymbol{Q}_2=\begin{pmatrix} \boldsymbol{R}_1^{-1}\boldsymbol{B}_1\boldsymbol{R}_1 & & 0 \\ & \ddots & \\ 0 & & \boldsymbol{R}_t^{-1}\boldsymbol{B}_t\boldsymbol{R}_t \end{pmatrix}$$

为对角矩阵．至于

$$\boldsymbol{Q}_2^{-1}\boldsymbol{Q}_1^{-1}\boldsymbol{A}\boldsymbol{Q}_1\boldsymbol{Q}_2=\mathrm{diag}\,\{\lambda_1\boldsymbol{E}_1, \lambda_2\boldsymbol{E}_2, \cdots, \lambda_t\boldsymbol{E}_t\}$$

是显然的．这说明了 $\boldsymbol{P}=\boldsymbol{Q}_1\boldsymbol{Q}_2$ 的确可使 $\boldsymbol{A}$、$\boldsymbol{B}$ 同时化为对角矩阵．即证明了 $k=2$ 时引理成立．

再归纳地假定引理对 $k-1$ 个矩阵是成立的．现在来考虑引理中的 k 个矩阵 $\boldsymbol{A}_1, \boldsymbol{A}_2, \cdots, \boldsymbol{A}_k$．因为 $\boldsymbol{A}_1$ 与对角矩阵相似，故有 $\boldsymbol{C}$ 使

$$\boldsymbol{C}^{-1}\boldsymbol{A}_1\boldsymbol{C}=\mathrm{diag}\,\{\lambda_1\boldsymbol{E}_1, \lambda_2\boldsymbol{E}_2, \cdots, \lambda_t\boldsymbol{E}_t\},$$

$\lambda_i\neq\lambda_j(i\neq j$ 时）. 由于 $\boldsymbol{C}^{-1}\boldsymbol{A}_i\boldsymbol{C}$ 与 $\boldsymbol{C}^{-1}\boldsymbol{A}_1\boldsymbol{C}$ 相乘可交换 $(i\neq 1)$, 故

$$\boldsymbol{C}^{-1}\boldsymbol{A}_i\boldsymbol{C}=\begin{pmatrix}\boldsymbol{A}_i^{(1)} & & & 0\\ & \boldsymbol{A}_i^{(2)} & & \\ & & \ddots & \\ 0 & & & \boldsymbol{A}_i^{(t)}\end{pmatrix},$$

$\boldsymbol{A}_i^{(j)}$ 与 $\boldsymbol{E}_j$ 是同级的. 从 $\boldsymbol{A}_i(i=2, \cdots, k)$ 相似于对角矩阵知 $\lambda\boldsymbol{E}-\boldsymbol{A}_i$ 的初等因子为 λ 之一次式, 因而 $\lambda\boldsymbol{E}_j-\boldsymbol{A}_i^{(j)}$ 的初等因子亦为 λ 之一次式, 故 $\boldsymbol{A}_i^{(j)}$ 与对角矩阵相似; 但从 $\boldsymbol{C}^{-1}\boldsymbol{A}_i\boldsymbol{C}$ 与 $\boldsymbol{C}^{-1}\boldsymbol{A}_l\boldsymbol{C}$ 相乘可交换$(i, l=2, \cdots, k)$, 得知

$$\boldsymbol{A}_i^{(j)}\boldsymbol{A}_l^{(j)}=\boldsymbol{A}_l^{(j)}\boldsymbol{A}_i^{(j)}\quad (j=1, 2, \cdots, t),$$

于是, 对每一 $j=1, 2, \cdots, t$, $\boldsymbol{A}_2^{(j)}$, $\boldsymbol{A}_3^{(j)}$, $\cdots$, $\boldsymbol{A}_k^{(j)}$ 这 $k-1$ 个矩阵适合引理中的条件, 故据归纳法的假定, 得知有一个满秩矩阵 $\boldsymbol{R}_j$ 使对 $i=2, 3, \cdots, k$ 都有 $\boldsymbol{R}_j^{-1}\boldsymbol{A}_i^{(j)}\boldsymbol{R}_j$ 为对角矩阵. 再作

$$\boldsymbol{R}=\begin{pmatrix}\boldsymbol{R}_1 & & & 0\\ & \boldsymbol{R}_2 & & \\ & & \ddots & \\ 0 & & & \boldsymbol{R}_t\end{pmatrix},$$

则 $\boldsymbol{R}$ 是满秩的, 且显然有

$$\boldsymbol{R}^{-1}\boldsymbol{C}^{-1}\boldsymbol{A}_i\boldsymbol{C}\boldsymbol{R}=\begin{pmatrix}\boldsymbol{R}_1^{-1}\boldsymbol{A}_i^{(1)}\boldsymbol{R}_1 & & & 0\\ & \boldsymbol{R}_2^{-1}\boldsymbol{A}_i^{(2)}\boldsymbol{R}_2 & & \\ & & \ddots & \\ 0 & & & \boldsymbol{R}_t^{-1}\boldsymbol{A}_i^{(t)}\boldsymbol{R}_t\end{pmatrix}$$

为对角矩阵, 即 $\boldsymbol{P}^{-1}\boldsymbol{A}_i\boldsymbol{P}$ 为对角矩阵 $(i=2, 3, \cdots, k;\ \boldsymbol{P}=\boldsymbol{C}\boldsymbol{R})$. 至于

$$
\begin{aligned}
\boldsymbol{P}^{-1}\boldsymbol{A}_1\boldsymbol{P} &= \boldsymbol{R}^{-1}\boldsymbol{C}^{-1}\boldsymbol{A}_1\boldsymbol{C}\boldsymbol{R} \\
&= \boldsymbol{R}^{-1}\cdot\begin{pmatrix} \lambda_1\boldsymbol{E}_1 & & & \\ & \lambda_2\boldsymbol{E}_2 & & \\ & & \ddots & \\ & & & \lambda_t\boldsymbol{E}_t \end{pmatrix}\cdot\boldsymbol{R} \\
&= \begin{pmatrix} \lambda_1\boldsymbol{E}_1 & & & \\ & \lambda_2\boldsymbol{E}_2 & & \\ & & \ddots & \\ & & & \lambda_t\boldsymbol{E}_t \end{pmatrix}
\end{aligned}
$$

为对角阵是显见的. 于是引理完全获证.

练习题答案和提示

练习一 1. $x=-a$, $y=b$, $z=c$. 2. $(-1)^{\frac{n(n-1)}{2}}$. 3. 因非零的元素少于 n，而每项又必有 n 个元，故为零。

练习二 1. 将第一列、第二列分别乘以 100 及 10 后再加到第三列上. 2. 由下一行减上面的行，可知结果等于零. 3. 利用行列式的性质 3，可将 D_n 分解成 2^n 个 n 级行列式之和，再利用行列式的性质 2 的推论 2，得知 $D_n=0(n\geqslant 3)$，但 $D_2=(x_2-x_1)(y_2-2y_1)$. 4. 将行列式之每元 a_{ij} 代以 1 即知. 5. 考虑转置行列式 D_n'，可知 D_n' 之每行为 D_n 之行的反号，故 $(-1)^nD_n=D_n'$.

练习三 1. 经行的互换，(i) 可化为范得蒙行列式，结果是

$$(-1)^{\frac{n(n+1)}{2}}\prod_{0\leqslant j<i\leqslant n}(i-j)=(-1)^{\frac{n(n+1)}{2}}n!\,(n-1)!\cdots\cdots 2!;$$

又，从 (ii) 之第 i 行提出公因子 $a_i^n\ (i=1,\ 2,\ \cdots,\ n+1)$ 得结果

$$(a_1a_2\cdots a_na_{n+1})^n\prod_{1\leqslant j<i\leqslant n+1}\left(\frac{b_i}{a_i}-\frac{b_j}{a_j}\right)=\prod_{1\leqslant j<i\leqslant n+1}(b_ia_j-b_ja_i).$$

2. (i) $D_n=(a+b)D_{n-1}-abD_{n-2}$，而 $D_1=a+b$，$D_2=(a+b)^2-ab=a^2+ab+b^2$;

(ii) $D_n=2\cos\theta\cdot D_{n-1}-D_{n-2}=2\cos\theta\dfrac{\sin n\theta}{\sin\theta}-\dfrac{\sin(n-1)\theta}{\sin\theta}=\dfrac{\sin(n+1)\theta}{\sin\theta}$.

3. 用 $x,\ x^2,\ \cdots,\ x^{n-2},\ x^{n-1}$ 分别去乘行列式之第二、三…，$n-1$，n 列，然后统统加到第一列上，再依第一列展开. 4. 关系式为 $D_n=aD_{n-1}-b\cdot cD_{n-2}$. 在 $a=1+bc$ 时易证 $D_n=1+bc+b^2c^2+\cdots+b^nc^n$. 5. 将所有的列都加到第一列上，易看出行列式有因子 x，即原行列式为

$$x\begin{vmatrix}1 & a_2 & a_3 & \cdots & a_n\\ 1 & x-\Delta_2 & a_3 & \cdots & a_n\\ \vdots & \vdots & \vdots & & \vdots\\ 1 & a_2 & a_3 & \cdots & x-\Delta_n\end{vmatrix}=x\begin{vmatrix}1 & a_2 & a_3 & \cdots & a_n\\ 0 & x-s & 0 & \cdots & 0\\ 0 & 0 & x-s & \cdots & 0\\ \vdots & \vdots & \vdots & & \vdots\\ 0 & 0 & 0 & \cdots & x-s\end{vmatrix}=x(x-s)^{n-1}.$$

6. (i) 令行列式 D 之第 2、3、…、n 列分别提出公因子 2, 3, …, n 后，得一个 n 级范得蒙行列式，故 $D=n!\sum\limits_{n\geqslant i>j\geqslant 1}(i-j)=n!\,(n-1)!\,(n-2)!\cdots 3!\,2!\,1!$;

(ii) 利用第 2 节中的性质可将行列式 D 分解为 2^n 个 n 级行列式之和，但其中凡有两列同为由数 x 所组成的显然等于零，故结果可知

$$D=\begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix}+\sum_{i=1}^{n} xD_i,$$

而 D_i 是将 $\begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \cdots & \cdots & \cdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix}$ 之第 i 列换成数 1 所成之 n 级行列式，因之 $D_i=A_{1i}+A_{2i}+\cdots+A_{ni}$，于是有

$$D=A+\Big(\sum_{i,j=1}^{n} A_{ij}\Big)x,$$

A_{ij} 为 $A=\begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \cdots & \cdots & \cdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix}$ 中元 a_{ij} 之代数余子式；

(iii) 由 D 之各行减去第 1 行，得

$$D=\begin{vmatrix} 1 & n & n & \cdots & n & n \\ n-1 & -(n-2) & 0 & \cdots & 0 & 0 \\ n-1 & 0 & -(n-3) & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ n-1 & 0 & 0 & \cdots & -1 & 0 \\ n-1 & 0 & 0 & \cdots & 0 & 0 \end{vmatrix}$$

$$=(-1)^{n+1}(n-1)\cdot(-1)^{n-1+1}n\cdot(-1)^{n-2}(n-2)!$$

$$=(-1)^{n-1}\cdot n!;$$

(iv) 由第二行减第一行，再由第三行减第二行，…，一直到由第 n 行减第 $n-1$ 行，则得原行列式

$$D=\begin{vmatrix} 1 & 2 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 1 & 2 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & 2 & \cdots & 0 & 0 \\ 0 & 0 & 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 1 & 2 \\ 0 & 0 & 0 & 0 & \cdots & 0 & 1 \end{vmatrix}=1;$$

(v) 令 Δ_n 表示给出的 n 级行列式，则 $\Delta_n=(1+x^2)\cdot\Delta_{n-1}-x^2\cdot\Delta_{n-2}$，故 $\Delta_n-\Delta_{n-1}=x^2\cdot(\Delta_{n-1}-\Delta_{n-2})$；但 $\Delta_1=1+x^2$，$\Delta_2=1+x^2+x^4$，故可递推得 $\Delta_n=1+x^2+x^4+\cdots+x^{2n}$.

练习四 1. 利用克莱姆定理得 $Dx=D_1$，再将 D_1 可分解为三个行列式之和，其中有二个为零，另一个为 aD，故 $x=a$. 同理，有 $y=b$，$z=c$. 3. 设 $f(x)=c_nx^n+c_{n-1}x^{n-1}+\cdots+c_1x+c_0$. 具性质 $f(a_i)=0(i=1, 2, \cdots, n+1)$ 且 $a_1, a_2, \cdots,$

α_n, α_{n+1} 两两互异，则得具 $n+1$ 个未知量 $c_0, c_1, \cdots, c_{n-1}, c_n$ 的 $n+1$ 个齐次线性方程的方程组，其系数之行列式为范得蒙行列式 $\prod\limits_{1\le j<i\le n+1}(\alpha_i-\alpha_j)\ne 0$，故必 $c_0=c_1=\cdots=c_{n-1}=c_n=0$.

练习五 1. D 先得展开为 n^n 个行列式之和，如

$$D=\sum\begin{vmatrix} a_{1i_1} & a_{1i_2} & \cdots & a_{1i_n} \\ a_{2i_1} & a_{2i_2} & \cdots & a_{2i_n} \\ \vdots & \vdots & & \vdots \\ a_{ni_1} & a_{ni_2} & \cdots & a_{ni_n}\end{vmatrix} b_{i_11}b_{i_22}\cdots b_{i_nn},$$

再去掉 $i_1, \cdots, i_n$ 中有相同的那些项，可得

$$D=\left(\sum(-1)^{[i_1,i_2,\cdots,i_n]}b_{i_11}b_{i_22}\cdots b_{i_nn}\right)\cdot D_1=D_1D_2.$$

2. $\Delta=(a+b+c+d)(a+b-c-d)(a-b+c-d)(a-b-c+d)$. 3. 按第一、二、六行利用拉普拉斯展开式，计算的结果为 -27. 4. 提示：可令 $\varphi_1=\alpha_2-\alpha_3$, $\varphi_2=\alpha_3-\alpha_1$, $\varphi_3=\alpha_1-\alpha_2$，于是原行列式为

$$\begin{vmatrix} \cos(\alpha_1-\alpha_1) & \cos(\alpha_1-\alpha_2) & \cos(\alpha_1-\alpha_3) \\ \cos(\alpha_2-\alpha_1) & \cos(\alpha_2-\alpha_2) & \cos(\alpha_2-\alpha_3) \\ \cos(\alpha_3-\alpha_1) & \cos(\alpha_3-\alpha_2) & \cos(\alpha_3-\alpha_3)\end{vmatrix}$$

$$=\begin{vmatrix} \cos\alpha_1 & \sin\alpha_1 & 0 \\ \cos\alpha_2 & \sin\alpha_2 & 0 \\ \cos\alpha_3 & \sin\alpha_3 & 0\end{vmatrix}\begin{vmatrix} \cos\alpha_1 & \cos\alpha_2 & \cos\alpha_3 \\ \sin\alpha_1 & \sin\alpha_2 & \sin\alpha_3 \\ 0 & 0 & 0\end{vmatrix}=0$$

当然也可以直接展开，例如以 $\cos\varphi_3=\cos(\varphi_1+\varphi_2)$ 代入之去计算. 5. 令左边 $2n$ 级行列式表以 D，当 $\mu\ne 0$ 时，则将 D 之第 1、2、…、n 列均乘以 μ，再将它的第 $n+1$、$n+2$、…、$2n$ 列都乘以 $-\lambda$ 后而分别加到第 1、2、…、n 列上，得

$$D=\frac{1}{\mu^n}\begin{vmatrix} a_{11}\mu-b_{11}\lambda & \cdots & a_{1n}\mu-b_{1n}\lambda & b_{11} & \cdots & b_{1n} \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{n1}\mu-b_{n1}\lambda & \cdots & a_{nn}\mu-b_{nn}\lambda & b_{n1} & \cdots & b_{nn} \\ 0 & \cdots & 0 & \mu & & \mathbf{0} \\ \vdots & & \vdots & & \ddots & \\ 0 & \cdots & 0 & \mathbf{0} & & \mu\end{vmatrix}$$

$$=\begin{vmatrix} a_{11}\mu-b_{11}\lambda & \cdots & a_{1n}\mu-b_{1n}\lambda \\ \vdots & & \vdots \\ a_{n1}\mu-b_{n1}\lambda & \cdots & a_{nn}\mu-b_{nn}\lambda\end{vmatrix}.$$

若 $\mu=0$，直接验证.

练习六 1. 不一定；线性无关. 2. 不一定. 3. 都不一定. 4. 逆不一定成立. 5. 可能有，可能有，没有. 6. 秩为 0 时选 λ_i 均为 0，秩为 1 时任两行线

性相关，不妨设$(a_{11}, \cdots, a_{1n})$不为$\mathbf{0}$，则

$$(a_{i1}, \cdots, a_{in})=\lambda_i(a_{11}, \cdots, a_{1n}), \quad (i=1, \cdots, m),$$

取$\mu_j=a_{1j}(j=1, \cdots, n)$。

练习七 **1.** $\lambda\neq 1$、-2时有唯一解；$\lambda=1$时有无穷多解；$\lambda=-2$时无解。 **2.** 如有解，易知$\sum a_i=0$；反之，若$\sum a_i=0$，则易知方程组的n个行向量线性相关，系数矩阵与增广矩阵的秩都是$n-1$，故有解。解前$n-1$个方程，得：

$$x_{n-1}=x_n+a_{n-1},\ x_{n-2}=x_n+a_{n-1}+a_{n-2},\ \cdots,\ x_2=x_n+a_{n-1}+\cdots+a_3+a_2,$$
$$x_1=x_n+a_{n-1}+\cdots+a_3+a_2+a_1.$$

3. 由假设易知系数矩阵与增广矩阵有等秩。 **4.** 设$\boldsymbol{A}$，$\boldsymbol{B}$分别为系数矩阵与增广矩阵。因$1\leqslant r_A\leqslant r_B\leqslant 3$，$r_B=r_A$或$r_B=r_A+1$，故有五种可能。(一)$r_B=r_A=3$时，三平面交一点；(二)$r_B=3$，$r_A=2$时，三平面中任二平面之交线与另一平面平行，或三平面中有二个平行，而另一个与它们各交一直线；(三)$r_B=r_A=2$时，三平面交于一直线(可能有二个平面重合)；(四)$r_B=2$，$r_A=1$时，三平面互平行(可能有二个重合)；(五)$r_B=r_A=1$时，三平面重合。**5.** 交一点时，应有

$$0=\begin{vmatrix} a & b & c\\ b & c & a\\ c & a & b\end{vmatrix}=-(a+b+c)[(a-b)^2+(b-c)^2+(c-a)^2],$$

故$a+b+c=0$；反之，$a+b+c=0$时，易证$\begin{pmatrix} a & b & c\\ b & c & a\\ c & a & b\end{pmatrix}$与$\begin{pmatrix} a & b\\ b & c\\ c & a\end{pmatrix}$之秩均为2。

练习八 **1.** 若$\sum\limits_{1}^{k}c_i(\boldsymbol{\alpha}_i+\boldsymbol{\alpha}_{i+1})=\mathbf{0}$(令$\boldsymbol{\alpha}_{k+1}=\boldsymbol{\alpha}_1$)，则

$$(c_k+c_1)\boldsymbol{\alpha}_1+(c_1+c_2)\boldsymbol{\alpha}_2+\cdots+(c_{k-1}+c_k)\boldsymbol{\alpha}_k=\mathbf{0},\ c_k=-c_1,$$

但又$c_k=-c_{k-1}=(-1)^2c_{k-2}=\cdots=(-1)^{k-1}c_1=c_1$，故$c_k=0$，因之$c_1=\cdots=c_{k-1}=0$。**2.** 原方程组与添加$b_1x_1+\cdots+b_nx_n=0$进去所得的方程组同解，因此它们系数矩阵之秩相等，故行向量$(b_1, b_2,\cdots,b_n)$为前面m个行向量$(a_{i1},a_{i2},\cdots,a_{in})$的线性组合$(i=1, \cdots, m)$。 **3.** 直接代入去检验。 **4.** 秩$=2$，每基础解系含有3个无关的解。 **5.** 若有不全为0的$\lambda_1, \lambda_2, \cdots, \lambda_s$使$\sum\limits_{i=1}^{s}\lambda_ix_i=0$，则令$|\lambda_p|=\max(|\lambda_1|, \cdots, |\lambda_s|)$，就有$\lambda_p\neq 0$，且由$\sum\limits_{i=1}^{s}\lambda_ix_i=0$之第$p$个分量可知

$$\sum_{i=1}^{s}\lambda_ia_{ip}=0,$$

故$\lambda_pa_{pp}=-\sum\limits_{\substack{i=1\\ i\neq p}}^{s}\lambda_ia_{ip}$，由之即得$|a_{pp}|\leqslant\sum\limits_{\substack{i=1\\ i\neq p}}^{s}\left|\dfrac{\lambda_i}{\lambda_p}\right||a_{ip}|\leqslant\sum\limits_{\substack{i=1\\ i\neq p}}^{s}|a_{ip}|$，矛盾了。**6.** 系数矩阵$(a_{ij})$之秩$r\leqslant n$，$(i=1, 2, \cdots, n; j=1, 2, \cdots, n+2)$，故只需解其中$r$个

方程，且可使 $n+2-r(\geqslant 2)$ 个未知量移至方程的右端作为常数，再利用克莱姆定理． **7**．移项后变成了 n 个未知量 $x_1,\cdots,x_n$ 的 n 个线性齐次方程的方程组，而系数行列式

$$\begin{vmatrix}1-n & 1 & 1 & \cdots & 1\\ 1 & 1-n & 1 & \cdots & 1\\ \vdots & \vdots & \vdots & & \vdots\\ 1 & 1 & 1 & \cdots & 1-n\end{vmatrix}=0,$$

但左上角的 $n-1$ 级主子式 $=(-1)^{n-1}n^{n-2}$，故其秩 $=n-1$，而得任意解为 $(\lambda,\lambda,\cdots,\lambda)$ 形．

练习九　1． 增广矩阵

$$\boldsymbol{B}=\begin{pmatrix}a & 1 & 1 & p\\ 1 & a & 1 & q\\ 1 & 1 & a & r\end{pmatrix}\sim\begin{pmatrix}a+2 & a+2 & a+2 & p+q+r\\ 1 & a & 1 & q\\ 1 & 1 & a & r\end{pmatrix};$$

当 $a+2\neq 0$ 时，

$$\boldsymbol{B}\simeq\begin{pmatrix}1 & 1 & 1 & \dfrac{p+q+r}{a+2}\\ 1 & a & 1 & q\\ 1 & 1 & a & r\end{pmatrix}\simeq\begin{pmatrix}1 & 1 & 1 & \dfrac{p+q+r}{a+2}\\ 0 & a-1 & 0 & \dfrac{aq-p+q-r}{a+2}\\ 0 & 0 & a-1 & \dfrac{(ar-p-q+r)}{a+2}\end{pmatrix};$$

又，当 $a-1\neq 0$ 时，有

$$\boldsymbol{B}\simeq\begin{pmatrix}1 & 1 & 1 & \dfrac{p+q+r}{a+2}\\ 0 & 1 & 0 & \dfrac{aq-p+q-r}{(a+2)(a-1)}\\ 0 & 0 & 1 & \dfrac{ar-p-q+r}{(a+2)(a-1)}\end{pmatrix}\simeq\begin{pmatrix}1 & 0 & 0 & \dfrac{ap+p-q-r}{(a+2)(a-1)}\\ 0 & 1 & 0 & \dfrac{aq-p+q-r}{(a+2)(a-1)}\\ 0 & 0 & 1 & \dfrac{ar-p-q+r}{(a+2)(a-1)}\end{pmatrix}.$$

于是，在 $(a+2)(a-1)\neq 0$ 时，方程组有唯一组解

$$x=\frac{ap+p-q-r}{(a+2)(a-1)},\quad y=\frac{aq-p+q-r}{(a+2)(a-1)},\quad z=\frac{ar-p-q+r}{(a+2)(a-1)}.$$

在 $a+2=0$，$p+q+r=0$ 时，

$$\boldsymbol{B}\simeq\begin{pmatrix}0 & 0 & 0 & 0\\ 1 & -2 & 1 & q\\ 1 & 1 & -2 & r\end{pmatrix}\simeq\begin{pmatrix}0 & 0 & 0 & 0\\ 1 & -2 & 1 & q\\ 0 & 3 & -3 & r-q\end{pmatrix}$$

$$\simeq\begin{pmatrix}0 & 0 & 0 & 0\\ 1 & -2 & 1 & q\\ 0 & 1 & -1 & \dfrac{r-q}{3}\end{pmatrix}\simeq\begin{pmatrix}0 & 0 & 0 & 0\\ 1 & 0 & -1 & \dfrac{2r+q}{3}\\ 0 & 1 & -1 & \dfrac{r-q}{3}\end{pmatrix},$$

这时方程组之解是 $x=\frac{2r+q}{3}+z$，$y=\frac{r-q}{3}+z$（z 任意）。在 $a=1$ 时，若 $2q=p+r$，$2r=p+q$，则 $x=p-y-z$（y，z 任意）。此外一切情况，方程组都无解。 **2**．分 2° 型及 3° 型两种行初等变换的情况，来证明(*)、(**)的对应分量的线性关系)。 **3**．先把第 i 行加到第 j 行上，再以 -1 乘第 j 行加到第 i 行，又把第 i 行加到第 j 行，最后以 -1 乘第 i 行。结果即为第 i，j 两行互换。

练习十　**1**．设 $R=(r_1, r_2, \cdots, r_n)$，$S=(s_1, s_2, \cdots, s_n)$，则 $r_i=x_i+\lambda(y_i-x_i)$，$s_i=x_i+\mu(y_i-x_i)$；又若 $Z=(z_1, z_2, \cdots, z_n)$ 使 $z_i=r_i+\sigma(s_i-r_i)$，则

$$z_i=x_i+[\lambda(1-\sigma)+\sigma\mu](y_i-x_i).$$

故当 λ、μ、σ 都在区间 $[0, 1]$ 内时，易证 $0\leqslant\lambda(1-\sigma)+\sigma\mu\leqslant 1$。

练习十一　**1**．选 $\boldsymbol{\alpha}_3$、$\boldsymbol{\alpha}_4$ 之方法很多，只要由 $\alpha_i(i=1, 2, 3, 4)$ 为行向量之行列式 $\neq 0$ 即可。 **2**．维数 $=2$。

练习十二　**1**．(1) y-轴($n=2$)，平面 YOZ($n=3$)；(2) 二、四象限之分角线($n=2$)，平面 $x+y=0$($n=3$)。 **2**．$\mathfrak{S}_1\cap\mathfrak{S}_2\neq 0$ 说明有 $\mathbf{0}\neq\boldsymbol{\alpha}\in\mathfrak{S}_1\cap\mathfrak{S}_2$，故 $\boldsymbol{\alpha}=\boldsymbol{\alpha}+\mathbf{0}=\mathbf{0}+\boldsymbol{\alpha}$ 之表示式非唯一。反之，若 $\mathfrak{S}_1\cap\mathfrak{S}_2=0$，则从 $\boldsymbol{\alpha}=\boldsymbol{\alpha}_1+\boldsymbol{\alpha}_2=\boldsymbol{\beta}_1+\boldsymbol{\beta}_1$ 得 $\boldsymbol{\alpha}_1-\boldsymbol{\beta}_1=\boldsymbol{\beta}_2-\boldsymbol{\alpha}_2\in\mathfrak{S}_1\cap\mathfrak{S}_2$，不得不有 $\boldsymbol{\alpha}_1=\boldsymbol{\beta}_1$，$\boldsymbol{\alpha}_2=\boldsymbol{\beta}_2$。 **3**．由 $\boldsymbol{\alpha}_1\in\mathfrak{S}_1$，$\boldsymbol{\alpha}_1\notin\mathfrak{S}_2$ 和 $\boldsymbol{\alpha}_2\notin\mathfrak{S}_1$，$\boldsymbol{\alpha}_2\in\mathfrak{S}_2$，即得 $\boldsymbol{\alpha}_1+\boldsymbol{\alpha}_2\notin\mathfrak{S}_i(i=1, 2)$。 **4**．设两个子空间 $\mathfrak{S}_1$，$\mathfrak{S}_2$ 之维数分别为 n_1，n_2，令 $\mathfrak{S}_1\cap\mathfrak{S}_2$ 之维数 $=d$，则 $\mathfrak{S}_1\cup\mathfrak{S}_2$ 之维数 $=d+1$，故再由 $d\leqslant n_i\leqslant d+1(i=1, 2)$ 知每 $n_i=d$ 或 $=d+1$，只此二个可能。又当 $n_1=n_2$ 时，若它们 $=d$，则必 $\mathfrak{S}_1=\mathfrak{S}_1\cap\mathfrak{S}_2=\mathfrak{S}_2$，因之 $\mathfrak{S}_1\cup\mathfrak{S}_2=\mathfrak{S}_1=\mathfrak{S}_2=\mathfrak{S}_1\cap\mathfrak{S}_2$，与 $\mathfrak{S}_1\cup\mathfrak{S}_2$ 之维数 $=d+1$ 相抵，故 $n_1\neq n_2$，于是 n_1，n_2 中必有一为 d，另一为 $d+1$。

练习十三　**1**．设 $\boldsymbol{a}_i=(a_{i1}, a_{i2}, \cdots, a_{in})$ 为 $\mathfrak{L}$ 之基底($i=1, 2, \cdots, k$)，考虑线性方程组 $\sum\limits_{j=1}^{n}a_{ij}x_j=0(i=1, 2, \cdots, k)$，作它的一组基础解系 $\boldsymbol{b}_i=(b_{i1}, b_{i2}, \cdots, b_{in})$，($i=1, 2, \cdots, n-k$)，易知 $\mathfrak{L}$ 即为方程组 $\sum\limits_{j=1}^{n}b_{ij}x_j=0$ 之解向量空间。 **2**．设 $\boldsymbol{ABX}=\mathbf{0}$ 之解空间为 $\mathfrak{N}$，$\boldsymbol{BX}=\mathbf{0}$ 之解空间为 $\mathfrak{M}$，则 $\mathfrak{M}\subseteq\mathfrak{N}$；又 $\dim\mathfrak{N}=n-r_{AB}$，$\dim\mathfrak{M}=n-r_B$，故 $r_{AB}=r_B$ 之充要条件是 $\mathfrak{M}=\mathfrak{N}$。 **3**．由上题可知 $\boldsymbol{BAX}=\mathbf{0}$ 必导致 $\boldsymbol{AX}=\mathbf{0}$，因而 $\boldsymbol{BACX}=\mathbf{0}$ 又必含有 $\boldsymbol{ACX}=\mathbf{0}$，故再由上题得 $r_{AC}=r_{BAC}$。 **4**．利用上题，将 $\boldsymbol{A}$、$\boldsymbol{B}$、$\boldsymbol{C}$ 分别代以 $\boldsymbol{A}^k$，$\boldsymbol{A}$，$\boldsymbol{A}$。

练习十四　**1**．$\sum\limits_{i=0}^{n-1}c_i(x-1)^i=0$ 可逐次地得到 $c_{n-1}=0$，$c_{n-2}=0$，$\cdots$，$c_1=0$，$c_0=0$。又 $f(x)=\sum\limits_{k=0}^{n-1}\frac{f^{(k)}(1)}{k!}\cdot(x-1)^k$。 **2**．$a\oplus x=b$ 说明 $ax=b$，$x=a^{-1}b$，即有“加法”的逆运算“减法”。定义 1 中条件 2) 易于检验。取任何正实数 μ 后，则对任正实数 a，欲 $a=k\odot\mu=\mu^k$ 成立之条件是 $k=\log_\mu a$。

练习十五　**1**．$\sum\limits_{1}^{4}k_i\boldsymbol{\beta}_i=\mathbf{0}$ 可顺次得到 $k_4=0$，$k_3=0$，$k_2=0$，$k_1=0$。所求之矩阵为

$$\begin{pmatrix} -1 & -1 & 1 & 2 \\ -2 & -4 & 3 & 3 \\ -1 & -8 & 9 & 2 \\ -1 & -9 & 9 & 4 \end{pmatrix}.$$ **2**. 1) 不是； 2) 不是， 3) 不是， 4) 是， 5) 不是。

3. $\begin{pmatrix} \cos^2\theta & \frac{1}{2}\sin 2\theta \\ \frac{1}{2}\sin 2\theta & \sin^2\theta \end{pmatrix}$，但 $\operatorname{tg}\theta=-\dfrac{a}{b}$。

练习十六 **1**. 设 $X=(x_{ij})$，则由 $DX=XD$ 可知 $c_ix_{ij}=c_jx_{ij}$，故 $i\neq j$ 时有 $x_{ij}=0$。 **3**. 条件均为 $AB=BA$。 **4**. $A=(a_{ij})$，$B=(b_{ij})\Rightarrow AB-BA$ 之主对角线上元的和为 $\sum\limits_{i,l=1}^{n} a_{il}b_{li}-b_{il}a_{li}=0$。 **5**. 由矩阵乘积之定义即得。

练习十八 **1**. 由定理 3 知 $A'\cdot(A^{-1})'=(A^{-1}A)'=E$，故 $(A')^{-1}=(A^{-1})'$。利用 $A^*=|A|\cdot A^{-1}$ 可知 $A^*\cdot(A^{-1})^*=|A|A^{-1}\cdot|A^{-1}|(A^{-1})^{-1}=|A|\,|A^{-1}|A^{-1}A=E$，即 $(A^{-1})^*=(A^*)^{-1}$。再直接利用伴随及转置之意义可证 $(A^*)'=(A')^*$。 **2**. 适合 $AX=E$ 之 X 存在的充要条件是 $\sum\limits_{l=1}^{n+1}a_{il}x_{lj}=\delta_{ij}(i,j=1,2,\cdots,n)$，于是，对每个 j，有 n 个方程 $\sum\limits_{l=1}^{n+1}a_{il}x_{lj}=\delta_{ij}(i=1,2,\cdots,n)$ 组成的线性方程组，有解之充要条件是 E 之每列为 A 之列的线性组合。 **3**. 仿上题，线性方程组

$$\sum_{l=1}^{n}a_{il}x_{lj}=b_{ij}\quad(i=1,2,\cdots,n)$$

有解之充要条件是 B 之第 j 列为 A 之 n 个列的线性组合。 **4**. 先将 X 写成分块之形。 **5**. 当 A 为满秩时，从 $AC=CA$ 得 $A^{-1}C=CA^{-1}$，因而由

$$\begin{pmatrix} E & 0 \\ -CA^{-1} & E \end{pmatrix}\begin{pmatrix} A & B \\ C & D \end{pmatrix}=\begin{pmatrix} A & B \\ 0 & D-CA^{-1}B \end{pmatrix}$$

而取行列式可知

$$\begin{vmatrix} A & B \\ C & D \end{vmatrix}=|A|\cdot|D-CA^{-1}B|=|A(D-CA^{-1}B)|=|AD-CB|.$$

当 A 为降秩时，考虑 $A+xE$（x 为未定元），因必有一常数 k 使 $x\geqslant k$ 时，$A+xE$ 恒为满秩的，于是由于 $(A+xE)C=C(A+xE)$，而据上述可知

$$\begin{vmatrix} A+xE & B \\ C & D \end{vmatrix}=|AD+xD-CB|,$$

由于这是关于 x 的恒等式，故以 $x=0$ 代入之，得

$$\begin{vmatrix} A & B \\ C & D \end{vmatrix}=|AD-CB|.$$

练习十九 **1**. $A\simeq E_0$（E_0 之意义参见第 177 页），故 $E_0=C_{11}+C_{22}+\cdots+C_{rr}$（$C_{ij}$ 之意义参见第 140 页）；再从 $A=PE_0Q=PC_{11}Q+\cdots+PC_{rr}Q$ 以及

$r_{PC_{ij}Q}=1$ 即得; **2**. 设 $\boldsymbol{A}=(a_{ij})$, $\boldsymbol{B}=(b_{ij})$, a_{ij} 与 b_{ij} 在 $\boldsymbol{A}$ 及 $\boldsymbol{B}$ 内的代数余子式分别为 $\boldsymbol{A}_{ij}$ 与 $\boldsymbol{B}_{ij}$, 则 $\boldsymbol{P}=\boldsymbol{AB}=\sum\limits_{t=1}^{n}a_{it}b_{tj}$, 故 $\boldsymbol{P}^*=(\boldsymbol{AB})^*$ 之第 j 行 i 列处的元 $=\sum a_{it}b_{tj}$ 之代数余子式 $=(-1)^{i+j}\cdot d(\boldsymbol{P}^{1,\cdots,i-1,i+1,\cdots,n}_{1,\cdots,j-1,j+1,\cdots,n})$

$$
\begin{aligned}
&=(-1)^{i+j}\sum d(\boldsymbol{A}^{1,\cdots,i-1,i+1,\cdots,n}_{h_1,\cdots,h_{n-1}})\cdot d(\boldsymbol{B}^{h_1,\cdots,h_{n-1}}_{1,\cdots,j-1,j+1,\cdots,n})\\
&=\sum(-1)^{i+h_m}d(\boldsymbol{A}^{1,\cdots,i-1,i+1,\cdots,n}_{h_1,h_2,\cdots,h_{n-1}})\\
&\quad\cdot(-1)^{j+h_m}d(\boldsymbol{B}^{h_1,h_2,\cdots,h_{n-1}}_{1,\cdots,j-1,j+1,\cdots,n})\\
&=\sum \boldsymbol{A}_{it}\boldsymbol{B}_{tj}=\boldsymbol{B}^*
\end{aligned}
$$

之第 j 行与 $\boldsymbol{A}^*$ 之第 i 列各相应元之积之和, 即 $(\boldsymbol{AB})^*=\boldsymbol{B}^*\boldsymbol{A}^*$. 即证得了 $k=2$ 时的正确性. 然后, 再用归纳法, 推广到 k 为多个的情形.

练习二十 **1**. 初等因子为 $\lambda-\lambda_1$ 与 $\lambda-\lambda_2$, 不变因子为 1 与 $(\lambda-\lambda_1)(\lambda-\lambda_2)$, 标准形为 $\begin{pmatrix}1 & 0\\ 0 & (\lambda-\lambda_1)(\lambda-\lambda_2)\end{pmatrix}$. **2**. 初等因子为 $\lambda,\lambda,\lambda^2,\lambda-1,(\lambda-1)^2,(\lambda-1)^3$; 不变因子因而为 1, $\lambda(\lambda-1)$, $\lambda(\lambda-1)^2$, $\lambda^2(\lambda-1)^3$. **3**. $D^{\boldsymbol{A}}_{k-1}(\lambda)\cdot D^{\boldsymbol{A}}_{k+1}(\lambda)=D^{\boldsymbol{A}}_{k-1}(\lambda)\cdot D^{\boldsymbol{A}}_{k}(\lambda)\cdot I^{\boldsymbol{A}}_{k+1}(\lambda)=D^{\boldsymbol{A}}_{k}(\lambda)\cdot D^{\boldsymbol{A}}_{k-1}(\lambda)\cdot I^{\boldsymbol{A}}_{k}(\lambda)\cdot\varphi(\lambda)=[D^{\boldsymbol{A}}_{k}(\lambda)]^2\cdot\varphi(\lambda)$.
4. 先将 $\boldsymbol{A}(\lambda)$ 之第 4、5、6 列都乘以 $\lambda-a$, 再分别加到第 1、2、3 列上; 再将所得之新矩阵的第 1、2、3 行都用 $\lambda-a$ 去乘后分别加到第 4、5、6 行上; 最后又将所得矩阵之第 4、5、6 列与第 1、2、3 列分别互换, 即得 $\boldsymbol{A}(\lambda)\simeq\boldsymbol{B}(\lambda)$. 再利用定理 6, 易知 $\boldsymbol{A}(\lambda)$ 之初等因子与三级 λ-矩阵

$$
\boldsymbol{P}(\lambda)=\begin{pmatrix}(\lambda-a)^2+b^2 & 1 & 0\\ 0 & (\lambda-a)^2+b^2 & 1\\ 0 & 0 & (\lambda-a)^2+b^2\end{pmatrix}
$$

的一致. 又, 易知 $D^{\boldsymbol{P}}_2(\lambda)=1$, 故 $D^{\boldsymbol{P}}_1(\lambda)=1$, 但

$$
D^{\boldsymbol{P}}_3(\lambda)=[(\lambda-a)^2+b^2]^3=[\lambda-(a+bi)]^3[\lambda-(a-bi)]^3,
$$

故 $I^{\boldsymbol{P}}_1(\lambda)=I^{\boldsymbol{P}}_2(\lambda)=1$, $I^{\boldsymbol{P}}_3(\lambda)=D^{\boldsymbol{P}}_3(\lambda)$, 因之 $\boldsymbol{P}(\lambda)$ 的初等因子为 $[\lambda-(a+bi)]^3$ 与 $[\lambda-(a-bi)]^3$, 于是

$$
\boldsymbol{A}(\lambda)\simeq\begin{pmatrix}1 &&&&&\\ &1&&&&\\ &&1&&&\\ &&&1&&\\ &&&&1&\\ &&&&&[(\lambda-a)^2+b^2]^3\end{pmatrix},
$$

当 $b\neq0$ 时. 至于 $b=0$ 时, $I^{\boldsymbol{A}}_6(\lambda)=(\lambda-a)^6$, 标准形仍为上形(令 $b=0$), 但初等因子只有 $(\lambda-a)^6$ 一个. **5**. 删去 $\lambda\boldsymbol{E}-\boldsymbol{A}$ 之第一列及末行, 得一个 $2n-1$ 级子式 $\neq0$, 而 $|\lambda\boldsymbol{E}-\boldsymbol{A}|=[(\lambda-a)^2+b^2]^n$, 故不变因子为 $\underbrace{1,1,\cdots,1}_{2n-1\text{个}}$, $[(\lambda-a)^2$

$+b^2]^n$；初等因子仅为$(\lambda-a+ib)^n$及$(\lambda-a-ib)^n$；约旦标准形为

$$\begin{pmatrix} a-ib & 1 & & & & & & & & & \\ & a-ib & 1 & & & & & & & & \\ & & a-ib & 1 & & & & & & & \\ & & & \ddots & \ddots & & & & & & \\ & & & & a-ib & 1 & & & & & \\ & & & & & a-ib & 1 & & & & \\ & & & & & & a+ib & 1 & & & \\ & & & & & & & a+ib & 1 & & \\ & & & & & & & & a+ib & 1 & \\ & & & & & & & & & \ddots & \ddots \\ & & & & & & & & & a+ib & 1 \\ & & & & & & & & & & a+ib \end{pmatrix}.$$

练习二十一　1．$\lambda\boldsymbol{E}-\boldsymbol{A}$之$D_1^A(\lambda)=1$，$D_2^A(\lambda)=(\lambda-1)(\lambda-2)$；对$\lambda\boldsymbol{E}-\boldsymbol{B}$也如此，可令

$$\boldsymbol{M}(\lambda)=\begin{pmatrix} 1 & 0 \\ \dfrac{3(\lambda-1)}{2} & 1 \end{pmatrix},\quad \boldsymbol{N}(\lambda)=\begin{pmatrix} -2 & 0 \\ \dfrac{-3(\lambda-2)}{2} & -\dfrac{1}{2} \end{pmatrix}.$$

但注意$\boldsymbol{M}(\lambda)$与$\boldsymbol{N}(\lambda)$的非唯一．　2．与证第3节定理1完全相仿．

练习二十二　1．(i) $m_A(\lambda)=\varphi(\lambda)\cdot q(\lambda)\Rightarrow 0=m_A(\boldsymbol{A})=\varphi(\boldsymbol{A})\cdot q(\boldsymbol{A})$；(ii) $d(\lambda)=p(\lambda)\varphi(\lambda)+q(\lambda)\cdot m_A(\lambda)\Rightarrow d(\boldsymbol{A})=p(\boldsymbol{A})\cdot\varphi(\boldsymbol{A})\Rightarrow r_{d(A)}\leqslant r_{\varphi(A)}$，又

$$\varphi(\lambda)=f(\lambda)\cdot d(\lambda)\Rightarrow\varphi(\boldsymbol{A})=f(\boldsymbol{A})\cdot d(\boldsymbol{A})\Rightarrow r_{\varphi(A)}\leqslant r_{d(A)};$$

(iii) $r_{\varphi(A)}=n\Rightarrow r_{d(A)}=n$（由(ii)），故由(i)知$d(\lambda)$为非0之常数，即$\varphi(\lambda)$与$m_A(\lambda)$互质；反之，应有

$$\varphi(\lambda)\cdot q(\lambda)+p(\lambda)\cdot m_A(\lambda)=1\Rightarrow\varphi_{(A)}\cdot q_{(A)}=\boldsymbol{E};$$

2．$|\lambda\boldsymbol{E}-\boldsymbol{A}|=\Delta_A(\lambda)=\lambda^3-\lambda^2-\lambda+1$，故$\boldsymbol{A}^n=\boldsymbol{A}^{n-2}+\boldsymbol{A}^2-\boldsymbol{E}$在$n=3$时成立，再用归纳法知对任$n>3$时亦成立，因而

$$\boldsymbol{A}^{100}=\boldsymbol{A}^{98}+(\boldsymbol{A}^2-\boldsymbol{E})=\boldsymbol{A}^{96}+2(\boldsymbol{A}^2-\boldsymbol{E})=\cdots=\boldsymbol{A}^2+49\cdot(\boldsymbol{A}^2-\boldsymbol{E})$$

$$=\begin{pmatrix} 1 & 0 & 0 \\ 50 & 1 & 0 \\ 50 & 0 & 1 \end{pmatrix}.$$

3．$\boldsymbol{A}$之特征值为$\lambda_1,\cdots,\lambda_n\Rightarrow\boldsymbol{A}^2$的为$\lambda_i^2\ (i=1,\cdots,n)\Rightarrow\lambda_i^2=1\Rightarrow\lambda_i=\pm1$．若每$\lambda_i=1$，则$\boldsymbol{A}+\boldsymbol{E}$满秩，因之$\boldsymbol{O}=(\boldsymbol{A}+\boldsymbol{E})(\boldsymbol{A}-\boldsymbol{E})\Rightarrow\boldsymbol{A}=\boldsymbol{E}$．　4．$\varphi(\lambda)=\lambda^k-1$无重根及$\varphi(A)=0\Rightarrow m_A(\lambda)$无重根．　5．$\boldsymbol{A}^2$之特征值为

$$\lambda_1^2,\cdots,\lambda_n^2\Rightarrow\sum_{i=1}^n\lambda_i^2=\sum_{i=1}^n\sum_{l=1}^n a_{il}a_{li}.$$

6. 由 $AX=\lambda X\Rightarrow A^kX=\lambda^kX$(用归纳法)$\Rightarrow f(A)X=f(\lambda)X$. 7. $\operatorname{tr}(AX)=0$ $\Rightarrow\sum_{i,j=1}^{n}a_{ij}x_{ji}=0\Rightarrow$令 $x_{st}=1$ 而其它 $x_{ij}=0$ 时就有 $a_{ts}=0$. 8. $A^2=nA\Rightarrow m_A(\lambda)=\lambda^2-n\lambda$, 故$A$之特征值只能是 0 与 n. 但 $r_A=1\Rightarrow A$ 之特征值有 $n-1$ 个 0 及仅一个 n. 9. $r_A=2\Rightarrow\Delta_A(\lambda)=\lambda^n-\sum_{i=1}^{n}(2i-1)\cdot\lambda^{n-1}+$ (A 中一切 2 级主子式之和)$\cdot\lambda^{n-2}=\lambda^n-n^2\cdot\lambda^{n-1}+\lambda^{n-2}\cdot\alpha$, 但

$$\begin{aligned}-\alpha&=\left[\sum_1^{n-1}i^2+\sum_1^{n-2}i^2+\sum_1^{n-3}i^2+\cdots+1^2\right]\\&=\frac{1}{6}[(n-1)n(2n-1)+(n-2)(n-1)(2n-3)+\cdots+1\cdot2\cdot3]\\&=\frac{n^2(n-1)(n+1)}{12},\end{aligned}$$

故

$$\Delta_A(\lambda)=\lambda^{n-2}\left[\lambda^2-n^2\lambda-\frac{1}{12}(n-1)n^2(n+1)\right].$$

10. $x(\lambda)=|\lambda E-A|=\lambda^2-5\lambda+7$, 而

$$\lambda^4-5\lambda^3+6\lambda^2+6\lambda-8=x(\lambda)\cdot(\lambda^2-1)+(\lambda-1),$$

故由 $x(A)=0$ 得

$$\begin{aligned}(A^4-5A^3+6A^2+6A-8E)^{-1}&=(A-E)^{-1}\\&=\begin{pmatrix}1&-1\\1&2\end{pmatrix}^{-1}=\frac{1}{3}\begin{pmatrix}2&1\\-1&1\end{pmatrix}.\end{aligned}$$

11. A 之特征多项式为 $x(\lambda)=|\lambda E-A|=(\lambda+1)(\lambda-1)(\lambda-2)$, 故令

$$\lambda^{2n}=x(\lambda)\cdot q(\lambda)+a\lambda^2+b\lambda+c$$

时, 应有

$$\begin{cases}1=a+b+c,\\1=a-b+c,\\2^{2n}=4a+2b+c,\end{cases}$$

解之得 $b=0$, $a=\frac{1}{3}(2^{2n}-1)$, $c=-\frac{1}{3}(2^{2n}-4)$, 于是

$$A^{2n}=\frac{1}{3}\{(2^{2n}-1)A^2-(2^{2n}-4)E\}.$$

12. $\sum\omega_i^k=\operatorname{tr}A^k=\sum a_{i\lambda_1}a_{\lambda_1\lambda_2}\cdots a_{\lambda_{k-1}i}$. 13. 有满秩矩阵 P、Q, 使

$$P^{-1}AQ^{-1}=\begin{pmatrix}E_r&O\\O&O\end{pmatrix},$$

$A=P\begin{pmatrix}E_r&0\\0&0\end{pmatrix}Q=\begin{pmatrix}P_1&P_2\\P_3&P_4\end{pmatrix}\begin{pmatrix}E_r&0\\0&0\end{pmatrix}\begin{pmatrix}Q_1&Q_2\\Q_3&Q_4\end{pmatrix}=\begin{pmatrix}P_1E_r&0\\P_3E_r&0\end{pmatrix}\begin{pmatrix}Q_1&Q_2\\Q_3&Q_4\end{pmatrix}=\begin{pmatrix}P_1E_rQ_1&P_1E_rQ_2\\ & \end{pmatrix}$

再令

$$P=\begin{pmatrix} P_1 & P_2 \\ \underbrace{P_3}_{r} & P_4 \end{pmatrix}\}r,\quad Q=\begin{pmatrix} Q_1 & Q_2 \\ \underbrace{Q_3}_{r} & Q_4 \end{pmatrix}\}r,$$

可知

$$\begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix} QP \begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix}=\begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix},$$

故 $Q_1P_1+Q_2P_3=E_r$，因而又有 $r=\mathrm{tr}(Q_1P_1)+\mathrm{tr}(Q_2P_3)$，又因恒有 $\mathrm{tr}(AB)=\mathrm{tr}(BA)$，故 $\mathrm{tr}\,A=r$.

练习二十三 1. $\varphi(\lambda)=\lambda^k-1$ 无重根，且 $m_A(\lambda)\mid\varphi(\lambda)$. 2. $\varphi(A)=O$, $\varphi(\lambda)=\lambda^k$，故 $m_A(\lambda)=\lambda^l(2\leqslant l\leqslant k)$. 3. $\varphi(\lambda)=\lambda^2+\lambda-2$ 无重根，故 $m_A(\lambda)$ 也无重根. 4. $\Delta_A(\lambda)=|\lambda E-A|$ 有二个不同的实根. 5. $\varphi(\lambda)=\lambda^2-\lambda$ 无重根，$m_A(\lambda)\mid\varphi(\lambda)$，故 $m_A(\lambda)$ 之根只能为 0 与 1，且 A 与对角矩阵相似，又从

$$P^{-1}AP=\begin{pmatrix} E_r & O \\ O & O \end{pmatrix}$$

得

$$P^{-1}(A+E)P=\begin{pmatrix} 2E_r & O \\ O & E_{n-r} \end{pmatrix},$$

故 $|A+E|=2^r$. 6. $|\lambda E-A|$ 之四个根为 i, $-i$, w, w^2($i=\sqrt{-1}$, w 为 1 之一虚立方根)，故 $P^{-1}AP=\mathrm{diag}\{i,\ -i,\ w,\ w^2\}$, $A^{60}=E$. 7. 有 P 使 $P^{-1}AP=\mathrm{diag}\{\lambda_1,\ \lambda_2,\ \cdots,\ \lambda_n\}$，并有 $P^{-1}AP\cdot P^{-1}BP=P^{-1}BP\cdot P^{-1}AP$，故由于 $\lambda_1,\cdots,\lambda_n$ 两两互异，因之 $P^{-1}BP=\mathrm{diag}\{\mu_1,\ \mu_2,\ \cdots,\ \mu_n\}$(参看练习十六第 1 题)，再利用拉格朗日内插法，知有唯一个 $n-1$ 次的多项式 $f(x)$，使 $f(\lambda_i)=\mu_i$ $(i=1,\ 2,\ \cdots,\ n)$. 于是

$$P^{-1}\cdot f(A)\cdot P=\mathrm{diag}\{f(\lambda_1),\ f(\lambda_2),\ \cdots,\ f(\lambda_n)\}=\mathrm{diag}\{\mu_1,\ \mu_2,\ \cdots,\ \mu_n\}.$$

8. 因有满秩矩阵 P、Q，使

$$PAQ=\begin{pmatrix} E_r & O \\ O & O \end{pmatrix}=C(\text{令}),$$

则 $C^2=C$. 于是 $A=P^{-1}CQ^{-1}=P^{-1}Q^{-1}C_1$, $C_1=QCQ^{-1}$, $C_1^2=C_1$; 或

$$A=C_2P^{-1}Q^{-1},\ C_2=P^{-1}CP,\ C_2^2=C_2.$$

9. A 与 B 均相似于同一个对角矩阵，故有满秩阵 Q 使 $Q^{-1}AQ=B$，因之 $Q^{-1}A=BQ^{-1}=R$.

练习二十四 1. A 之幂零性 $\Rightarrow A$ 之 n 个特征值 $\lambda_1=\cdots=\lambda_n=0$，故约旦标准形为 $\begin{pmatrix} 0 & & \mathbf{0} \\ & 0 \ddots & \\ * & & 0 \end{pmatrix}$，因而 $A+E\sim\begin{pmatrix} 1 & & \mathbf{0} \\ & 1 \ddots & \\ * & & 1 \end{pmatrix}$，故 $|A+E|=1$. 2. 由练习十六

之第 5 题得知 n 级约旦矩阵 $\boldsymbol{J}=\sum_{i=1}^{n-1}\boldsymbol{C}_{i+1,i}$，故易证 $\boldsymbol{J}^k=\sum_{i=1}^{n-k}\boldsymbol{C}_{i+k,i}$，于是 $\boldsymbol{J}^n=0$，因 $\boldsymbol{A}\sim\begin{pmatrix}\boldsymbol{J}_1 & & \\ & \ddots & \\ & & \boldsymbol{J}_t\end{pmatrix}$，故 $\boldsymbol{A}^r=0$； **3**．取 $1, x, x^2, \cdots, x^{n-1}$ 这一组基底言，因 $\boldsymbol{D}x^{i-1}=(i-1)x^{i-2}$，故

$$\boldsymbol{D}=\begin{pmatrix}0 & 0 & 0 & \cdots & 0 & 0\\ 1 & 0 & 0 & \cdots & 0 & 0\\ 0 & 2 & 0 & \cdots & 0 & 0\\ \vdots & \vdots & \vdots & & \vdots & \vdots\\ 0 & 0 & 0 & \cdots & 0 & 0\\ 0 & 0 & 0 & \cdots & n-1 & 0\end{pmatrix},$$

因之 $|\lambda\boldsymbol{E}-\boldsymbol{D}|=\lambda^n$ 为所求特征多项式．由 $\boldsymbol{D}$ 之幂零性即得 $\boldsymbol{D}$ 不与对角矩阵相似． **4**．容易看出

$$\lambda\boldsymbol{E}-\boldsymbol{A}=\begin{pmatrix}\lambda & -1 & -1\\ -1 & \lambda & -1\\ -1 & -1 & \lambda\end{pmatrix}$$

中 $D_1^{\boldsymbol{A}}=1$，$D_2^{\boldsymbol{A}}(\lambda)=\lambda+1$，$D_3^{\boldsymbol{A}}(\lambda)=(\lambda+1)^2(\lambda-2)$，故

$$I_1^{\boldsymbol{A}}(\lambda)=1,\ I_2^{\boldsymbol{A}}(\lambda)=\lambda+1,\ I_3^{\boldsymbol{A}}(\lambda)=\lambda^2-\lambda-2,$$

因而 $\boldsymbol{A}$ 之有理标准形为 $\begin{pmatrix}-1 & & \\ & 0 & 2\\ & 1 & 1\end{pmatrix}$． **5**．易知 $\lambda\boldsymbol{E}-\boldsymbol{A}$ 中

$$D_1^{\boldsymbol{A}}(\lambda)=D_2^{\boldsymbol{A}}(\lambda)=1,\ D_3^{\boldsymbol{A}}(\lambda)=(\lambda-1)^3,$$

由是 $\lambda\boldsymbol{E}-\boldsymbol{A}$ 只有唯一个初等因子 $(\lambda-1)^3$，约旦标准形为 $\begin{pmatrix}1 & 0 & 0\\ 1 & 1 & 0\\ 0 & 1 & 1\end{pmatrix}$．

练习二十五 **1**．提示：$\varphi(x_1,x_2,x_3)=2x_1x_2+4x_1x_3+2x_2x_3$，$\begin{pmatrix}x_1\\ x_2\\ x_3\end{pmatrix}=\boldsymbol{Q}_1\begin{pmatrix}y_1\\ y_2\\ y_3\end{pmatrix}$，

$\boldsymbol{Q}_1=\begin{pmatrix}1 & -1 & 0\\ 1 & 1 & 0\\ 0 & 0 & 1\end{pmatrix}\Rightarrow\varphi=2y_1^2-2y_2^2+6y_1y_3-2y_2y_3=\boldsymbol{Y}'\boldsymbol{B}\boldsymbol{Y}$，$\boldsymbol{B}=\begin{pmatrix}2 & 0 & 3\\ 0 & -2 & -1\\ 3 & -1 & 0\end{pmatrix}$

$=\boldsymbol{Q}_1'\boldsymbol{A}\boldsymbol{Q}_1$． $\begin{pmatrix}y_1\\ y_2\\ y_3\end{pmatrix}=\boldsymbol{Q}_2\begin{pmatrix}z_1\\ z_2\\ z_3\end{pmatrix}$， $\boldsymbol{Q}_2=\begin{pmatrix}\frac{1}{2} & 0 & -\frac{3}{2}\\ 0 & 1 & 0\\ 0 & 0 & 1\end{pmatrix}\Rightarrow\varphi=\frac{1}{2}z_1^2-2z_2^2-$

$\frac{9}{2}z_3^2-2z_2z_3=Z'CZ$, $C=\begin{pmatrix}\frac{1}{2}&0&0\\0&-2&-1\\0&-1&-\frac{9}{2}\end{pmatrix}=Q_2'BQ_2$. 再令 $\begin{pmatrix}z_1\\z_2\\z_3\end{pmatrix}=Q_3\begin{pmatrix}t_1\\t_2\\t_3\end{pmatrix}$,

$Q_3=\begin{pmatrix}1&0&0\\0&-\frac{1}{2}&-\frac{1}{2}\\0&0&1\end{pmatrix}\Rightarrow\varphi=\frac{1}{2}t_1^2-\frac{1}{2}t_2^2-4t_3^2$ 为典型式, 变换矩阵为

$$Q=Q_1Q_2Q_3=\begin{pmatrix}\frac{1}{2}&\frac{1}{2}&-1\\\frac{1}{2}&-\frac{1}{2}&-2\\0&0&1\end{pmatrix}.$$

2. 提示: $AB=(AB)'=B'A'=BA$. **3**. 提示: $A=B+C$, $B=\frac{A+A'}{2}$, $C=\frac{A-A'}{2}$ 说明了存在性. 若又 $A=B_1+C_1(B_1'=B_1,\ C_1'=-C_1)$, 则 $B-B_1=C_1-C$, 左为对称, 右为反对称 $\Rightarrow(B-B_1)'=B-B_1=-(B-B_1)'\Rightarrow B-B_1=0$. **4**. 提示: $0=\begin{vmatrix}a_{ii}&a_{ij}\\a_{ji}&a_{jj}\end{vmatrix}\Rightarrow a_{ii}a_{jj}=a_{ij}^2$.

练习二十六 **1**. 提示: $P+N=r$ 与 $P-N=s$ 足以说明 r 与 s 同偶同奇, 且 $-r\leqslant s\leqslant r$. **2**. 提示: 由条件得知非负整数 $P=\frac{r+s}{2}$、$N=\frac{r-s}{2}$ 是存在的, 且 P、N 中至少有一个为正整数. **3**. 提示: 有满秩线性变换 $X=QZ$ 使 $\sum a_{ij}x_ix_j=z_1^2+\cdots+z_P^2-z_{P+1}^2-\cdots-z_{P+N}^2$, 再作满秩变换 $z_i=\sqrt{\alpha_i}y_i(i=1,\ \cdots,\ P)$, $z_{P+j}=\sqrt{\beta_j}y_{P+j}(j=1,\ \cdots,\ N)$, $z_{P+N+1}=y_{P+N+1}$, $\cdots$, $z_n=y_n$, 即得. **4**. 提示: 满秩变换 $X=A^{-1}Y$ 变

$$\sum a_{ij}x_ix_j=X'AX=Y'(A^{-1})'AA^{-1}Y=Y'(A^{-1})'Y=\sum\frac{A_{ij}}{a}y_iy_j.$$

5. 提示: 因恒有满秩变换 $X=Q_1Y$ 使 $\sum a_{ij}x_ix_j=\lambda_1y_1^2+\cdots+\lambda_ry_r^2$; 再作满秩变换

$$Y=Q_2Z,\quad Q_2=\begin{pmatrix}\frac{1}{\sqrt{\lambda_1}}&&&&&\\&\ddots&&&&\\&&\frac{1}{\sqrt{\lambda_r}}&&&\\&&&1&&\\&&&&\ddots&\\&&&&&1\end{pmatrix},$$

可得 $\sum a_{ij}x_ix_j=z_1^2+\cdots+z_r^2$; 再作 $z_i=\sqrt{\alpha_i}u_i(i=1,\ \cdots,\ r)$ 及 $z_j=u_j(j=r+1,\cdots,$

n)已使 $\sum a_{ij}x_ix_j=\alpha_1u_1^2+\cdots+\alpha_ru_r^2$. **6**. 提示: 设 $\sum a_{ij}x_ix_j=(\alpha_1x_1+\cdots+\alpha_nx_n)(\beta_1x_1+\cdots+\beta_nx_n)$, 由于 $\boldsymbol{A}=(a_{ij})\neq\boldsymbol{O}$, 知 α_1, …, α_n 不全为 0, 同理 β_1, …, β_n 不全为 0. 今设 $\alpha_1\neq0$, 并若 $\alpha_1x_1+\cdots+\alpha_nx_n$ 与 $\beta_1x_1+\cdots+\beta_nx_n$ 线性相关, 则有 ρ 使 $\beta_i=\rho\alpha_i(i=1, \cdots, n)$, 因之 $\rho\neq0$, $\sum a_{ij}x_ix_j=\rho(\alpha_1x_1+\cdots+\alpha_nx_n)^2$, 故满秩变换 $y_1=\alpha_1x_1+\cdots+\alpha_nx_n$, $y_i=x_i(i=2, \cdots, n)$ 已使 $\sum a_{ij}x_ix_j=\rho y_1^2$, 证明了 $r_{\boldsymbol{A}}=1$. 若 $\alpha_1x_1+\cdots+\alpha_nx_n$ 与 $\beta_1x_1+\cdots+\beta_nx_n$ 无关, 令 $\begin{vmatrix}\alpha_1 & \alpha_2\\ \beta_1 & \beta_2\end{vmatrix}\neq0$, 则满秩变换 $y_1=\sum\alpha_ix_i$, $y_2=\sum\beta_ix_i$, $y_j=x_j(j=3, \cdots, n)$ 使 $\sum a_{ij}x_ix_j=y_1y_2$, 而满秩变换 $y_1=z_1-z_2$, $y_2=z_1+z_2$, $y_j=z_j(j=3, \cdots, n)$ 使 $\sum a_{ij}x_ix_j=z_1^2-z_2^2$, $\therefore r_{\boldsymbol{A}}=2$. 反之, $r_{\boldsymbol{A}}=1\Rightarrow\sum a_{ij}x_ix_j=y_1^2$, $y_1=\alpha_1x_1+\cdots+\alpha_nx_n\Rightarrow\sum a_{ij}x_ix_j=(\sum\alpha_ix_i)^2$ 为二个相关一次因式之积. 若 $r_{\boldsymbol{A}}=2$, 典型式为 $\sum a_{ij}x_ix_j=y_1^2+y_2^2$, 满秩变换 $z_1=y_1+iy_2$, $z_2=y_1-iy_2$, $z_j=y_j(j=3, \cdots, n)$ 变 $\sum a_{ij}x_ix_j=z_1z_2$, 且易知 z_1 与 z_2 无关. **7**. 提示: 由下面的行减上面的行得

$$\boldsymbol{A}\simeq\begin{pmatrix}1 & 2 & 3 & 4\\ 1 & 1 & 1 & 1\\ 1 & 1 & 1 & 1\\ 1 & 1 & 1 & 1\end{pmatrix}\Rightarrow r_{\boldsymbol{A}}=2,$$

$$\boldsymbol{X}'\boldsymbol{A}\boldsymbol{X}=x_1^2+3x_2^2+5x_3^2+7x_4^2+4x_1x_2+\cdots\cdots$$

可由满秩变换 $y_1=x_1+2x_2+3x_3+4x_4$, $y_i=x_i(i=2, 3, 4)$ 使 $\boldsymbol{X}'\boldsymbol{A}\boldsymbol{X}=y_1^2+\psi$, 但 $\psi=-y_2^2-4y_3^2-9y_4^2-4y_2y_3-6y_2y_4-12y_3y_4=-z_2^2$, $z_2=-y_2-2y_3-3y_4$. 于是满秩变换 $z_1=y_1$, $z_2=-y_2-2y_3-3y_4$, $z_3=y_3$, $z_4=y_4$ 使 $\boldsymbol{X}'\boldsymbol{A}\boldsymbol{X}=z_1^2-z_2^2$, 符号差 $=0$, 故可分解, 且有 $\boldsymbol{X}'\boldsymbol{A}\boldsymbol{X}=(z_1-z_2)(z_1+z_2)=(x_1+3x_2+5x_3+7x_4)(x_1+x_2+x_3+x_4)$. 再归纳地假定 $n-1$ 级矩阵成立, 于是对 n 级 $\boldsymbol{A}$ 有

$$\begin{aligned}\boldsymbol{X}'\boldsymbol{A}\boldsymbol{X}&=[x_1+3x_2+5x_3+\cdots+(2n-3)x_{n-1}]\cdot(x_1+x_2+\cdots+x_{n-1})\\&\quad+2nx_n(x_1+x_2+\cdots+x_{n-1})+2x_n[x_2+2x_3+\cdots+(n-2)x_{n-1}]\\&\quad+(2n-1)x_n^2\\&=[x_1+3x_2+\cdots+(2n-3)x_{n-1}+(2n-1)x_n]\cdot(x_1+x_2+\cdots+x_{n-1})\\&\quad+x_n[x_1+3x_2+5x_3+\cdots+(2n-3)x_{n-1}]+(2n-1)x_n^2\\&=[x_1+3x_2+\cdots+(2n-3)x_{n-1}+(2n-1)x_n](x_1+x_2+\cdots+x_{n-1})\\&\quad+x_n[x_1+3x_2+\cdots+(2n-3)x_{n-1}+(2n-1)x_n]\\&=[x_1+3x_2+\cdots+(2n-1)x_n]\cdot(x_1+x_2+\cdots+x_n).\end{aligned}$$

8. 提示: 设 φ 之正负惯性指数分别为 k 与 l, 即有满秩线性变换 $y_i=a_{i1}x_1+\cdots+a_{in}x_n(i=1, \cdots, n)$ 使 $\boldsymbol{A}=(a_{ik})$ 是满秩的而有 $\varphi(x_1, \cdots, x_n)=y_1^2+\cdots+y_k^2-y_{k+1}^2-\cdots-y_{k+l}^2$. 如果 $k>p$, 则关于 n 个未知量 $x_1, \cdots, x_n$ 的线性方程组 $L_1=0, \cdots, L_p=0$, $y_{k+1}=0, \cdots, y_n=0$ 有非零解, 由是从 $\varphi(x_1, \cdots, x_n)=L_1^2+$

$\cdots+L_p^2-L_{p+1}^2-\cdots-L_{p+q}^2$ 得知 $\varphi(x_1, \cdots, x_n)\leqslant 0$，而从

$$\varphi=\sum_{i=1}^{k} y_i^2-\sum_{j=1}^{l} y_{k+j}^2$$

又得 $\varphi(x_1, \cdots, x_n)>0$（因 $y_1, \cdots, y_k$ 中至少有一个不等于 0），显相矛盾．故必 $k\leqslant p$．同理 $l\leqslant q$．

练习二十七 1．提示：$r_A=1$ 时有二类为 x_1^2 与 $-x_1^2$，一般 $r_A=k$ 时有 $k+1$ 个类，故齐式之类数 $=2+3+\cdots+(n+1)=\dfrac{1}{2}n(n+3)$．方程之类数在 $r_A=1$ 时为 1，在 $r_A=2$ 时为 2，$r_A=3$ 时亦为 2，一般 $r_A=2k$ 时有 $k+1$ 个类，在 $r_A=2k+1$ 时亦有 $k+1$ 个类．于是 $n=2m$ 时，

$$\begin{aligned}\text{类之总数}&=1+2+2+3+3+\cdots+m+m+(m+1)\\&=2(1+2+3+\cdots+m)+m=m(m+1)+m=\frac{n(n+4)}{4};\end{aligned}$$

$n=2m+1$ 时，类之总数

$$\begin{aligned}&1+2+2+3+3+\cdots+(m+1)+(m+1)\\&=2[1+2+3+\cdots+m+(m+1)]-1\\&=(m+1)(m+2)-1=\frac{n^2+4n-1}{4}.\end{aligned}$$

2．提示：因典型系数可限制为 1，故齐式之类均为 1，不论 r_A 为 $1, 2, \cdots, n$ 之任何值，因之齐式之类的总数 $=n$，方程之类数亦为 n．

练习二十八 1．提示：$\sum a_{ij}\bar{x}_i x_j=\overline{X}'AX$ 中

$$A=\begin{pmatrix}0 & 1-i & 2-2i\\ 1+i & 0 & 1-i\\ 2+2i & 1+i & 0\end{pmatrix},$$

令
$$X=Q_1Y,\quad Q_1=\begin{pmatrix}1-i & -1 & 0\\ 1 & 1+i & 0\\ 0 & 0 & 1\end{pmatrix},$$

则
$$\overline{X}'AX=\overline{Y}'BY,\quad B=\overline{Q}_1'AQ_1=\begin{pmatrix}4 & 1+i & 5-i\\ 1-i & -4 & -2\\ 5+i & -2 & 0\end{pmatrix};$$

再令
$$z_1=4y_1+(1+i)y_2+(5-i)y_3,\quad z_2=y_2,\quad z_3=y_3,$$

则
$$Y=Q_2Z,\quad Q_2=\begin{pmatrix}\frac{1}{4} & -\frac{1}{4}(1+i) & -\frac{5}{4}+\frac{1}{4}i\\ 0 & 1 & 0\\ 0 & 0 & 1\end{pmatrix},$$

$$\overline{X}'AX=\overline{Z}'CZ,$$

$$C=\bar{Q}_2'BQ_2=\begin{pmatrix}\frac{1}{4} & 0 & 0\\ 0 & -\frac{9}{2} & -3+\frac{3}{2}i\\ 0 & -3-\frac{3}{2}i & -\frac{13}{2}\end{pmatrix};$$

又令 $u_1=z_1,\quad u_2=-\frac{9}{2}z_2-\left(3-\frac{3}{2}i\right)z_3,\quad u_3=z_3,$

$$Z=Q_3U,\quad U=\begin{pmatrix}u_1\\u_2\\u_3\end{pmatrix},\quad Q_3=\begin{pmatrix}1 & 0 & 0\\ 0 & -\frac{2}{9} & -\frac{2}{3}+\frac{1}{3}i\\ 0 & 0 & 1\end{pmatrix},$$

则 $$\bar{X}'AX=\bar{U}'DU,\ D=\bar{Q}_3'CQ_3=\begin{pmatrix}4 & 0 & 0\\ 0 & -\frac{2}{9} & 0\\ 0 & 0 & -4\end{pmatrix},$$

故变换矩阵为 $Q_1Q_2Q_3$. **2**. 提示: $(\overline{AB})'=\bar{B}'\bar{A}'=BA$. **3**. 提示: 令 $B=\frac{A+\bar{A}'}{2}$, $C=\frac{A-\bar{A}'}{2}$, 则 $\bar{B}'=B$, $\bar{C}'=-C$ 且 $A=B+C$. 若又 $A=B_1+C_1(\bar{B}_1'=B_1,\ \bar{C}_1'=-C_1)$, 则 $B_1-B=C-C_1$, 故两边取共轭转置后可得 $B_1-B=C_1-C$,因之 $C=C_1$, $B=B_1$. 又由 $a_{ij}=-\bar{a}_{ji}$ 得 $a_{jj}=-\bar{a}_{jj}$, 即 a_{jj} 为纯虚数或 0. **4**. 提示: 令 $B=iA$. $\bar{A}'=A\Rightarrow\bar{B}'=-i\bar{A}'=-B$; $\bar{A}'=-A\Rightarrow\bar{B}'=-i\bar{A}'=B$. 又若 $\bar{A}'=-A$, 则 $|\bar{A}'|=|-A|=(-1)^n|A|$, 即 $|\bar{A}|=(-1)^n\cdot|A|\Rightarrow|\bar{A}|=|A|$或$=-|A|$ 由 n 为偶或奇而定. 又 $(\overline{A^k})'=(\bar{A}')^k=(-A)^k=(-1)^k\cdot A^k$. **5**. 提示: $A=(a_{st})$, $\bar{A}'=A$, $r_A=1$, $a_{ss}\cdot a_{tt}\neq0\Rightarrow0=\begin{vmatrix}a_{ss} & a_{st}\\ a_{ts} & a_{tt}\end{vmatrix}=a_{ss}a_{tt}-a_{st}\bar{a}_{st}\Rightarrow a_{ss}\cdot a_{tt}=a_{st}\bar{a}_{st}>0$. 当 $\bar{A}'=-A$ 时, $0=a_{ss}a_{tt}+a_{st}\bar{a}_{st}$. **6**. 提示: $\bar{A}'=A$ 及 $|A|\neq0\Rightarrow(\overline{A^{-1}})'=(\bar{A}')^{-1}=A^{-1}$. 同样可讨论反厄米特. **7**. 提示: 参看练习二十六的第 4 题. **8**. 提示: 参看练习二十六的第 3 题.

练习二十九 **1**. 提示: 从行列式之判别方法即得. **2**. 提示: $\varphi(x_1,\cdots,x_n)=\bar{X}'AX$, A 之正定性$\Rightarrow\varphi(0,\cdots,0,x_n)=a_{nn}\bar{x}_nx_n>0$ 当 $x_n\neq0$ 时$\Rightarrow a_{nn}>0$, 又 $\varphi(0,\cdots,0,x_{n-1},x_n)=(\bar{x}_{n-1},\bar{x}_n)\cdot\begin{pmatrix}a_{n-1,n-1} & a_{n-1,n}\\ a_{n,n-1} & a_{nn}\end{pmatrix}\cdot\begin{pmatrix}x_{n-1}\\x_n\end{pmatrix}>0$ 当 $\begin{pmatrix}x_{n-1}\\x_n\end{pmatrix}\neq0$ 时 $\Rightarrow\begin{vmatrix}a_{n-1,n-1} & a_{n-1,n}\\ a_{n,n-1} & a_{nn}\end{vmatrix}>0$, 继续这样下去, 可知

$$\begin{vmatrix}a_{kk} & \cdots & a_{kn}\\ \vdots & & \vdots\\ a_{nk} & \cdots & a_{nn}\end{vmatrix}>0.$$

条件之必要性获证．再假定 $n-1$ 级矩阵是成立的(证充分性时用归纳法)，于是因 $n-1$ 级矩阵

$$B=\begin{pmatrix} a_{22} & \cdots & a_{2n} \\ \vdots & & \vdots \\ a_{n2} & \cdots & a_{nn} \end{pmatrix}$$

满足题述条件，故 B 正定 $\Rightarrow n-1$ 级满秩矩阵 Q_1 使 $\bar{Q}_1'BQ_1=E$，再令

$$Q_2=\begin{pmatrix} 1 & 0 \\ 0 & Q_1 \end{pmatrix} \Rightarrow \bar{Q}_2'AQ_2=\begin{pmatrix} a_{11} & SQ_1 \\ \bar{Q}_1'\bar{S}' & E \end{pmatrix}, \quad S=(a_{12}, \cdots, a_{1n}),$$

令 $SQ_1=(b_{12}, \cdots, b_{1n})$，则从 $X=Q_2Y$ 得

$$\begin{aligned}\bar{X}'AX&=\bar{Y}'\bar{Q}_2'AQ_2Y=a_{11}\bar{y}_1y_1+\bar{y}_2y_2+\cdots+\bar{y}_ny_n+(b_{12}y_2+\cdots+b_{1n}y_n)\bar{y}_1\\&\quad+(\bar{b}_{12}\bar{y}_2+\cdots+\bar{b}_{1n}\bar{y}_n)y_1\\&=c\bar{y}_1y_1+(\bar{y}_2+b_{12}\bar{y}_1)(y_2+\bar{b}_{12}y_1)+\cdots+(\bar{y}_n+b_{1n}\bar{y}_1)(y_n+\bar{b}_{1n}y_1),\end{aligned}$$

$$c=a_{11}-b_{12}\bar{b}_{12}-\cdots-b_{1n}\bar{b}_{1n}$$

为实数，故满秩变换 $y_1=z_1$, $y_i+\bar{b}_{1i}y_1=z_i(i=2, \cdots, n)$，使

$$\bar{X}'AX=\bar{Y}'\bar{Q}_2'AQ_2Y=\bar{Z}'\bar{P}'\bar{Q}_2'AQ_2PZ=c\bar{z}_1z_1+\bar{z}_2z_2+\cdots+\bar{z}_nz_n,$$

即 $\bar{Q}'AQ=\operatorname{diag}\{c, 1, \cdots, 1\}$, $Q=Q_2P$，因之 $|A|\cdot|Q|\cdot|\bar{Q}|=c>0$，故 A 正定． **3.** 提示：若有 X 使 $\bar{X}'AX=-c$, $c>0$，则 $X=\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ 中 $x_1\neq 0$ (利用题设的条件及第 2 题)，由是 $\tau=c/\bar{x}_1x>0$，而 $\bar{X}'AX+\tau\bar{x}_1x_1=0$，此即说有 $X\neq 0$ 使

$$\bar{X}'BX=0, \quad B=\begin{pmatrix} a_{11}+\tau & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}.$$

另方面，

$$|B|=|A|+\tau\cdot\begin{vmatrix} a_{22} & \cdots & a_{2n} \\ \vdots & & \vdots \\ a_{n2} & \cdots & a_{nn} \end{vmatrix}=\tau\cdot\begin{vmatrix} a_{22} & \cdots & a_{2n} \\ \vdots & & \vdots \\ a_{n2} & \cdots & a_{nn} \end{vmatrix}>0,$$

故再由题设而据第 2 题知 B 正定．矛盾的产生即说明 A 为半正定的． **4.** 提示：利用行列式判别法即得． **5.** 提示：因存在 Q, $\bar{Q}'AQ=\operatorname{diag}\{a_1, \cdots, a_n\}$, $\bar{Q}'BQ=\operatorname{diag}\{b_1, \cdots, b_n\}$，故 $\bar{Q}'(A-\lambda B)Q=\operatorname{diag}\{a_1-\lambda b_1, \cdots, a_n-\lambda b_n\}$, $\Rightarrow|A-\lambda B|=0$ 之 n 个根为 $a_i/b_i(i=1, \cdots, n)$．再令

$$A=\begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}, \quad B=\begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix},$$

B 满秩，而 $|A-\lambda B|=0$ 之二根为 $\frac{1\pm\sqrt{3}i}{2}$，非实数，故 x^2+2xy 与 $2xy+y^2$ 不

能够借一个满秩变换同时变成平方和． 6．提示：存在满秩矩阵 $\boldsymbol{Q}$，$\overline{\boldsymbol{Q}}'\boldsymbol{A}\boldsymbol{Q}=\boldsymbol{E}\Rightarrow\boldsymbol{E}=(\overline{\boldsymbol{Q}}'\boldsymbol{A}\boldsymbol{Q})^{-1}=\boldsymbol{Q}^{-1}\cdot\boldsymbol{A}^{-1}\cdot(\overline{\boldsymbol{Q}^{-1}})'\Rightarrow\boldsymbol{A}^{-1}$ 正定． 又 $\boldsymbol{E}=\boldsymbol{E}^*=(\overline{\boldsymbol{Q}}'\boldsymbol{A}\boldsymbol{Q})^*=\boldsymbol{Q}^*\boldsymbol{A}^*(\overline{\boldsymbol{Q}^*})'=\overline{\boldsymbol{R}}'\boldsymbol{A}^*\boldsymbol{R}(\overline{\boldsymbol{R}}'=\boldsymbol{Q}^*)\Rightarrow\boldsymbol{A}^*$ 正定． 7．提示：

$$0<|\boldsymbol{A}|=\begin{vmatrix}\boldsymbol{A}_{n-1} & \boldsymbol{X}\\ \overline{\boldsymbol{X}}' & a_{nn}\end{vmatrix}=\begin{vmatrix}\boldsymbol{A}_{n-1} & \boldsymbol{X}\\ \overline{\boldsymbol{X}}' & 0\end{vmatrix}+\begin{vmatrix}\boldsymbol{A}_{n-1} & \boldsymbol{0}\\ \overline{\boldsymbol{X}}' & a_{nn}\end{vmatrix}$$
$$=-\overline{\boldsymbol{X}}'\boldsymbol{A}_{n-1}^*\boldsymbol{X}+a_{nn}|\boldsymbol{A}_{n-1}|.$$

8．提示：$\varphi(P)>0$ 与 $\varphi(Q)<0\Rightarrow P$，$Q$ 线性无关．若 $V(x_1,\cdots,x_n)$ 与 P、Q 线性相关，即 $x_i=\lambda y_i+\mu z_i(i=1,\cdots,n)$，使

$$\varphi(V)=0\Rightarrow 0=\lambda^2\cdot\sum a_{ij}y_iy_j+2\lambda\mu\cdot\sum a_{ij}y_iz_j+\mu^2\cdot\sum a_{ij}z_iz_j;$$

由于判别式 $4\{(\sum a_{ij}y_iz_j)^2-\varphi(P)\cdot\varphi(Q)\}>0$，知有 λ'/μ' 与 λ''/μ'' 互不等且 $x_i'=\lambda'y_i+\mu'z_i$ 及 $x_i''=\lambda''y_i+\mu''z_i$，$R(x_1'\cdots x_n')$ 及 $S(x_1''\cdots x_n'')$，$\varphi(R)=0=\varphi(S)$．
9．提示：原式$=\sum\limits_{i\neq j}(x_i-x_j)^2$． 10．提示：

$$原式=(x,\ y,\ z,\ t)\begin{pmatrix}\lambda & 1 & 1 & 0\\ 1 & \lambda & -1 & 0\\ 1 & -1 & \lambda & 0\\ 0 & 0 & 0 & 1\end{pmatrix}\begin{pmatrix}x\\ y\\ z\\ t\end{pmatrix},$$

故为正定之充要条件是 $\lambda>0$，$\lambda^2-1>0$，$(\lambda+1)^2(\lambda-2)>0$，因而为$\lambda>2$．当 $\lambda=2$ 时，原式$=\dfrac{1}{2}(2x+y+z)^2+\dfrac{3}{2}(y-z)^2+t^2=\dfrac{1}{2}u^2+\dfrac{3}{2}v^2+t^2$ 为半正定．

11．提示：$n=1$ 时显然成立．归纳地假定 $n-1$ 级时成立，由于

$$|\boldsymbol{A}|=a_{nn}\cdot|\boldsymbol{A}_{n-1}|-\overline{\boldsymbol{X}}'\boldsymbol{A}_{n-1}^*\boldsymbol{X},\quad \boldsymbol{X}=\begin{pmatrix}a_{1n}\\ \vdots\\ a_{n-1,n}\end{pmatrix}$$

（第 7 题），且 $\boldsymbol{A}_{n-1}^*$ 正定（第 6 题），故 $|\boldsymbol{A}|\leqslant a_{nn}|\boldsymbol{A}_{n-1}|$ 且“=”成立之充要条件是 $a_{1n}=\cdots=a_{n-1,n}=0$，由之得 $|\boldsymbol{A}|\leqslant a_{11}a_{22}\cdots a_{nn}$． 12．提示：$\boldsymbol{X}\neq\boldsymbol{0}$ 及 $\boldsymbol{Q}$ 之满秩性，即得 $\boldsymbol{Y}=\boldsymbol{Q}\boldsymbol{X}\neq\boldsymbol{0}$，因之 $\overline{\boldsymbol{X}}'(\overline{\boldsymbol{Q}}'\boldsymbol{A}\boldsymbol{Q})\boldsymbol{X}=(\overline{\boldsymbol{Q}\boldsymbol{X}})'\boldsymbol{A}(\boldsymbol{Q}\boldsymbol{X})=\overline{\boldsymbol{Y}}'\boldsymbol{A}\boldsymbol{Y}>0$，证明了 $\overline{\boldsymbol{Q}}'\boldsymbol{A}\boldsymbol{Q}$ 之正定性．对其它情形可类似地讨论． 13．提示：设 φ 为厄米特齐式（若为实二次齐式，同样可证明）．则

$$\begin{aligned}\varphi(x_1,\cdots,x_n)&=a_{11}x_1\bar{x}_1+\sum_{r=2}^{n}(a_{1r}x_1\bar{x}_r+a_{r1}x_r\bar{x}_1)+\sum_{r,s=2}^{n}a_{rs}x_r\bar{x}_s\\&=a_{11}\left\{\left(x_1+\frac{a_{21}}{a_{11}}x_2+\cdots+\frac{a_{n1}}{a_{11}}x_n\right)\left(\bar{x}_1+\frac{a_{12}}{a_{11}}\bar{x}_2+\cdots+\frac{a_{1n}}{a_{11}}\bar{x}_n\right)\right.\\&\qquad\left.-\sum_{r,s=2}^{n}\frac{a_{r1}}{a_{11}}x_r\frac{a_{1s}}{a_{11}}\bar{x}_s\right\}+\sum_{r,s=2}^{n}a_{rs}x_r\bar{x}_s\\&=a_{11}\left|x_1+\frac{a_{21}}{a_{11}}x_2+\cdots+\frac{a_{n1}}{a_{11}}x_n\right|^2+\sum_{r,s=2}^{n}\left(a_{rs}-\frac{a_{r1}a_{1s}}{a_{11}}\right)x_r\bar{x}_s;\end{aligned}$$

再令 $c_1=a_{11}>0,\quad \alpha_i=\dfrac{a_{i1}}{a_{11}},\quad y_1=x_1+\alpha_2x_2+\cdots+\alpha_nx_n,$

及 $$f(x_2,\ \cdots,\ x_n)=\sum_{r,s=2}^{n}\left(a_{rs}-\frac{a_{r1}a_{1s}}{a_{11}}\right)x_r\bar{x}_s,$$

则有 $\varphi(x_1,\ \cdots,\ x_n)=c_1y_1\bar{y}_1+f(x_2,\ \cdots,\ x_n)$，但 $f(x_2,\ \cdots,\ x_n)$ 是 $n-1$ 个变量 $x_2,\ \cdots,\ x_n$ 的厄米特齐式．如 $f(x_2,\ \cdots,\ x_n)\leqslant 0$，则令 $x_1=-(\alpha_2x_2+\cdots+\alpha_nx_n)$ 时就有 $\varphi(x_1,\ \cdots,\ x_n)=f(x_2,\ \cdots,\ x_n)\leqslant 0$，不可．故 $f(x_2,\ \cdots,\ x_n)$ 是 $n-1$ 个变量的正定厄米特齐式，再利用归纳法． **14**．提示：令 $\boldsymbol{S}$ 之秩$=r$，相异根之个数$=t$，而令 $\lambda_1,\ \lambda_2,\ \cdots,\ \lambda_t$ 为两两互异的，再作矩阵

$$\boldsymbol{A}=\begin{pmatrix}1 & 1 & \cdots & 1\\ \lambda_1 & \lambda_2 & \cdots & \lambda_n\\ \vdots & \vdots & & \vdots\\ \lambda_1^{n-1} & \lambda_2^{n-1} & \cdots & \lambda_n^{n-1}\end{pmatrix},$$

则 $\boldsymbol{S}=\boldsymbol{A}\boldsymbol{A}'$．显然，$\boldsymbol{A}$ 之左上角 t 级主子式$\neq 0$，且 $\boldsymbol{A}$ 之秩$=t$，故 $r\leqslant t$．另方面，设

$$\boldsymbol{M}=\begin{pmatrix}1 & 1 & \cdots & 1\\ \lambda_1 & \lambda_2 & \cdots & \lambda_n\\ \vdots & \vdots & & \vdots\\ \lambda_1^{t-1} & \lambda_2^{t-1} & \cdots & \lambda_n^{t-1}\end{pmatrix},$$

则 $$\boldsymbol{M}'=\begin{pmatrix}1 & \lambda_1 & \cdots & \lambda_1^{t-1}\\ 1 & \lambda_2 & \cdots & \lambda_2^{t-1}\\ \vdots & \vdots & & \vdots\\ 1 & \lambda_n & \cdots & \lambda_n^{t-1}\end{pmatrix},$$

而 $$\boldsymbol{M}\boldsymbol{M}'=\begin{pmatrix}s_0 & s_1 & \cdots & s_{t-1}\\ s_1 & s_2 & \cdots & s_t\\ \vdots & \vdots & & \vdots\\ s_{t-1} & s_t & \cdots & s_{2t-2}\end{pmatrix}$$

为 $\boldsymbol{S}$ 之左上角的 t 级子矩阵，故

$$\det \boldsymbol{M}\boldsymbol{M}'=\sum_{1\leqslant r_1<\cdots<r_t\leqslant n}\det\left(\boldsymbol{M}_{r_1r_2\cdots r_t}^{12\cdots t}\right)\cdot\det\left(\boldsymbol{M}'^{r_1r_2\cdots r_t}_{12\cdots t}\right)$$

$$=\sum_{1\leqslant r_1<\cdots<r_t\leqslant n}\begin{vmatrix}1 & 1 & \cdots & 1\\ \lambda_{r_1} & \lambda_{r_2} & \cdots & \lambda_{r_t}\\ \vdots & \vdots & & \vdots\\ \lambda_{r_1}^{t-1} & \lambda_{r_2}^{t-1} & \cdots & \lambda_{r_t}^{t-1}\end{vmatrix}^2$$

$$=m_1m_2\cdots m_t\begin{vmatrix}1 & 1 & \cdots & 1\\ \lambda_1 & \lambda_2 & \cdots & \lambda_t\\ \vdots & \vdots & & \vdots\\ \lambda_1^{t-1} & \lambda_2^{t-1} & \cdots & \lambda_t^{t-1}\end{vmatrix}^2\neq 0$$

(m_i 的 λ_i 的重复度)，又说明了 $t\leqslant r$． **15**．提示：$\boldsymbol{S}$ 为实对称矩阵．$\boldsymbol{S}=\boldsymbol{A}\boldsymbol{A}'$，

$\boldsymbol{A}$ 为满秩，$\boldsymbol{S}$ 因而为正定的．反之，设 $\boldsymbol{S}$ 已为正定的．若 $x^n+a_1x^{n-1}+\cdots+a_{n-1}x+a_n=0$ 有虚根 λ_1，则共轭虚数 $\bar{\lambda}_1$(设为 λ_2)亦必为根；故由 $|\boldsymbol{A}|\neq 0$，则知线性方程组

$$\begin{cases}x_1+\lambda_1x_2+\cdots+\lambda_1^{n-1}x_n=1\\ x_1+\lambda_2x_2+\cdots+\lambda_2^{n-1}x_n=-1\\ x_1+\lambda_ix_2+\cdots+\lambda_i^{n-1}x_n=0\quad(i=3,\ 4,\ \cdots,\ n)\end{cases}$$

有解 $\boldsymbol{X}'=(x_1,\ x_2,\ \cdots,\ x_n)$，故

$$\boldsymbol{X}'\boldsymbol{S}\overline{\boldsymbol{X}}=(x_1,\ \cdots,\ x_n)\boldsymbol{A}\boldsymbol{A}'\begin{pmatrix}\bar{x}_1\\ \vdots\\ \bar{x}_n\end{pmatrix}=(1,\ -1,\ 0,\ \cdots,\ 0)\begin{pmatrix}-1\\ 1\\ 0\\ \vdots\\ 0\end{pmatrix}=-2,$$

与 $\boldsymbol{S}$ 之正定相矛盾．故方程无虚根．　　**16**．提示：$\boldsymbol{A}$ 之任一特征值 $\lambda_i(i=1,2,\cdots,n)$ 具性质 $a\leqslant\lambda_i\leqslant b$ 时，则 $\boldsymbol{B}=\boldsymbol{A}-\lambda_0\boldsymbol{E}$ 之特征值 λ 应有 $0=|\lambda\boldsymbol{E}-\boldsymbol{B}|=|(\lambda+\lambda_0)\boldsymbol{E}-\boldsymbol{A}|\Rightarrow a\leqslant\lambda+\lambda_0\leqslant b\Rightarrow a-\lambda_0\leqslant\lambda\leqslant b-\lambda_0$，故 $\lambda_0<a$ 时 λ 是正数，而在 $\lambda_0>b$ 时 λ 为负数，即 $\boldsymbol{A}-\lambda_0\boldsymbol{E}$ 在 $\lambda_0<a$ 时为正定的，而在 $\lambda_0>b$ 时为负定的．反之亦然．　　**17**．提示：若 $\lambda_0<a+c$，则可记 $\lambda_0=\lambda_1+\lambda_2$ 使 $\lambda_1<a$ 及 $\lambda_2<c$，因而 $(\boldsymbol{A}+\boldsymbol{B})-\lambda_0\boldsymbol{E}=(\boldsymbol{A}-\lambda_1\boldsymbol{E})+(\boldsymbol{B}-\lambda_2\boldsymbol{E})$ 为正定的(利用上题)；又当 $\lambda_0>b+d$ 时可记 $\lambda_0=\mu_1+\mu_2$ 使 $\mu_1>b,\mu_2>d$，于是 $(\boldsymbol{A}+\boldsymbol{B})-\lambda_0\boldsymbol{E}=(\boldsymbol{A}-\mu_1\boldsymbol{E})+(\boldsymbol{B}-\mu_2\boldsymbol{E})$ 为负定的(仍利用上题)．故再度利用上题得 $\boldsymbol{A}+\boldsymbol{B}$ 之特征值全在区间 $[a+c,\ b+d]$ 内．

练习三十　**1**．提示：$\boldsymbol{A}$ 之正定性⇒使 $\mathfrak{R}$ 为欧氏空间，故 $\boldsymbol{x}=(x_1,\ \cdots,\ x_n)$，$\boldsymbol{y}=(y_1,\cdots,y_n)$，就有 $(\sum a_{ij}x_iy_j)^2=(\boldsymbol{x},\ \boldsymbol{y})^2\leqslant(\boldsymbol{x},\ \boldsymbol{x})\cdot(\boldsymbol{y},\boldsymbol{y})=(\sum a_{ij}x_ix_j)(\sum a_{ij}y_iy_j)$．**2**．提示：条件的必要性已在证定理 1 之推论时证了．今反之，设 $|(\boldsymbol{\alpha}_i,\ \boldsymbol{\alpha}_j)|=0$，则必有某列(例如最后一列)为它列的线性组合，因之存在 $\lambda_1,\ \cdots,\ \lambda_{k-1}$，使 $(\boldsymbol{\alpha}_i,\ \boldsymbol{\alpha}_k)=\lambda_1(\boldsymbol{\alpha}_i,\ \boldsymbol{\alpha}_1)+\cdots+\lambda_{k-1}(\boldsymbol{\alpha}_i,\ \boldsymbol{\alpha}_{k-1})=(\boldsymbol{\alpha}_i,\ \lambda_1\boldsymbol{\alpha}_1+\cdots+\lambda_{k-1}\boldsymbol{\alpha}_{k-1})$，$(i=1,2,\cdots,k)$，由是 $\left(\boldsymbol{\alpha}_i,\ \boldsymbol{\alpha}_k-\sum_1^{k-1}\lambda_i\boldsymbol{\alpha}_i\right)=0\Rightarrow\left(\boldsymbol{\alpha}_k-\sum_1^{k-1}\lambda_i\boldsymbol{\alpha}_i,\ \boldsymbol{\alpha}_k-\sum_1^{k-1}\lambda_i\boldsymbol{\alpha}_i\right)=0\Rightarrow\boldsymbol{\alpha}_k-\sum_1^{k-1}\lambda_i\boldsymbol{\alpha}_i=\boldsymbol{0}\Rightarrow\boldsymbol{\alpha}_1,\ \cdots,\ \boldsymbol{\alpha}_k$ 线性相关．　　**3**．提示：设 $M=(z_1,\ \cdots,\ z_n)$．因 M 在线段 PQ 上，故 $z_i=x_i+\lambda(y_i-x_i)$，$0\leqslant\lambda\leqslant 1$；因 $\overrightarrow{PM}=(z_1-x_1,\ \cdots,\ z_n-x_n)$，故 $\|\overrightarrow{PM}\|^2=\sum(z_i-x_i)^2=\sum\lambda^2(y_i-x_i)^2$，同理，$\|\overrightarrow{QM}\|^2=\sum(z_i-y_i)^2=\sum[x_i+\lambda(y_i-x_i)-y_i]^2=\sum(\lambda-1)^2(y_i-x_i)^2$．故从 $\|\overrightarrow{PM}\|=\|\overrightarrow{QM}\|$ 得 $0=\sum(y_i-x_i)^2[\lambda^2-(\lambda-1)^2]\Rightarrow 2\lambda-1=0\Rightarrow z_i=\dfrac{1}{2}(x_i+y_i)$，此即 i)．又若 $N=(z_1,\ \cdots,\ z_n)$ 且 $\|\overrightarrow{PN}\|=\|\overrightarrow{QN}\|$，则 $\sum(z_i-x_i)^2=\sum(z_i-y_i)^2\Rightarrow\sum 2(y_i-x_i)z_i=\sum(y_i^2-x_i^2)$，即超平面方程为 $\sum 2(y_i-x_i)X_i=\sum(y_i^2-x_i^2)$，而 $(X_1,\ \cdots,\ X_n)$ 为超平面上动点之坐标，此即

ii)．若 $R=(u_1, \cdots, u_n)$ 与 $S=(v_1, \cdots, v_n)$ 在这超平面上，则 $(\overrightarrow{PQ}, \overrightarrow{RS})=\sum(y_i-x_i)(v_i-u_i)$，再利用 $\sum 2(y_i-x_i)v_i=\sum(y_i^2-x_i^2)=\sum 2(y_i-x_i)u_i$ 得

$$\sum 2(y_i-x_i)(v_i-u_i)=0,$$

故证明了 iii)。

练习三十一 1．提示：从题设知 $e_j=\delta_{j1}e_1+\cdots+\delta_{jn}e_n$ 在 e_i 上之坐标

$$\delta_{ji}=\|e_j\|\cdot\cos(\widehat{e_j, e_i})=\|e_j\|\cdot\frac{(e_j, e_i)}{\|e_j\|\cdot\|e_i\|}=\frac{(e_j, e_i)}{\|e_i\|},$$

故 $i\neq j$ 时有 $(e_i, e_j)=0$，而

$$1=\delta_{ii}=\frac{(e_i, e_i)}{\|e_i\|}=\|e_i\|.$$

2．提示：

$$e_i=\sum_{t=1}^{n}\delta_{it}e_t,\ e_j=\sum_{t=1}^{n}\delta_{jt}e_t\Rightarrow(e_i, e_j)=\sum_{t=1}^{n}\delta_{it}\delta_{jt}=\delta_{ij}\delta_{jj}=\delta_{ij}.$$

3．提示：

$$(x_i, x_j)=\sum_{t=1}^{n}\xi_{it}\xi_{jt}\Rightarrow$$

$$\text{所求之格拉姆矩阵}=\begin{pmatrix}\sum\xi_{1t}\xi_{1t} & \cdots & \sum\xi_{1t}\xi_{kt}\\ \vdots & & \vdots\\ \sum\xi_{kt}\xi_{1t} & \cdots & \sum\xi_{kt}\xi_{kt}\end{pmatrix}$$

$$=\begin{pmatrix}\xi_{11} & \cdots & \xi_{1n}\\ \vdots & & \vdots\\ \xi_{k1} & \cdots & \xi_{kn}\end{pmatrix}\begin{pmatrix}\xi_{11} & \cdots & \xi_{k1}\\ \vdots & & \vdots\\ \xi_{1n} & \cdots & \xi_{kn}\end{pmatrix}=AA'.$$

4．提示：$x_i=\sum\xi_{it}e_t(i=1, \cdots, k)$ 及 $e_t(i=1, \cdots, n)$ 之标准直交性 $\Rightarrow(x_i, e_s)=\xi_{is}$，$(i=1, \cdots, k; s=1, \cdots, n)$；故 $(x_i, x_j)=\sum\xi_{it}\xi_{jt}=\sum(x_i, e_t)\cdot(x_j, e_t)$．

5．提示：易知 $B=\bar{A}'A=(b_{ij})$ 是正定的，故由练习二十九之第 11 题得知 $|A||\bar{A}|\leqslant b_{11}b_{22}\cdots b_{nn}$，但 $b_{kk}=\bar{a}_{1k}a_{1k}+\bar{a}_{2k}a_{2k}+\cdots+\bar{a}_{nk}a_{nk}$，且等号成立之条件是 $b_{ij}=0(i\neq j)$ 即 $a_{1i}\bar{a}_{1j}+\cdots+a_{ni}\bar{a}_{nj}=0(i\neq j)$．其几何意义实质上就是格拉姆行列式之几何意义．

练习三十二 1．提示：$(x, y)=(x\mathscr{T}, y\mathscr{T})\Rightarrow\|x\|\cdot\|y\|\cdot\cos(\widehat{x, y})=\|x\mathscr{T}\|\cdot\|y\mathscr{T}\|\cdot\cos(\widehat{x\mathscr{T}, y\mathscr{T}})$．　2．提示：$\|x-y\|^2=(x-y, x-y)=\|x\|^2+\|y\|^2-2\cdot(x, y)$，同理有 $\|x\mathscr{T}-y\mathscr{T}\|^2=\|x\mathscr{T}\|^2+\|y\mathscr{T}\|^2-2\cdot(x\mathscr{T}, y\mathscr{T})$．故由定义 1 及定理 1 即得．　3．提示：$P^{-1}AP$ 之直交性由定理 3 之推论可得．例 3 之 $A=\begin{pmatrix}0 & -1\\ -1 & 0\end{pmatrix}$ 直交，$\begin{pmatrix}e_1\\ e_2\end{pmatrix}=P\begin{pmatrix}i\\ j\end{pmatrix}$ 中之 $P=\begin{pmatrix}\frac{1}{2} & 1\\ -\frac{1}{2} & 1\end{pmatrix}$ 非直交，$P^{-1}AP$ 直交．　4．提示：$(A^{-1}SA)'=A'S'(A^{-1})'=A^{-1}S'A=\begin{cases}A^{-1}SA(S'=S\text{ 时}),\\ -A^{-1}SA(S'=-S\text{ 时}).\end{cases}$

5. 提示：利用定理 3 之 iv) 或 v)。 **6**. 提示：$A'=A$ 与 $A'=A^{-1}\Rightarrow A=A^{-1}$，即 $A^2=E$. 其它类推. **7**. 提示：因 $(a_{i1},\cdots,a_{in})$ 与 $(A_{i1},\cdots,A_{in})$ 均为 $n-1$ 个线性齐次方程 $a_{j1}x_1+\cdots+a_{jn}x_n=0$ 之非零解 ($j=1,\cdots,n$; 但 $j\neq i$)，其秩 $=n-1$，故它们成比例，即 $\dfrac{a_{i1}}{A_{i1}}=\cdots=\dfrac{a_{in}}{A_{in}}=\lambda$. 再利用 $\pm1=a_{i1}A_{i1}+\cdots+a_{in}A_{in}=\dfrac{1}{\lambda}(a_{i1}^2+\cdots+a_{in}^2)=\dfrac{1}{\lambda}$，即得. **8**. 提示：$n=1$ 时显然成立(这时只有两个向量所成的角大于 90° 而实际上等于 180°). 再归纳地假定 $k-1$ 维欧氏空间是成立的，而来考虑 k 维空间 $\mathfrak{R}_k$. 如若 $\mathfrak{R}_k$ 中存在着 $k+2$ 个向量 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_{k+2}$ 使得 $(\boldsymbol{\alpha}_i,\boldsymbol{\alpha}_j)$ 为负数 $(i\neq j)$. 今选取 $\mathfrak{R}_k$ 之一组标准直交基底 $\boldsymbol{e}_1\left(=\dfrac{1}{\|\boldsymbol{\alpha}_1\|}\boldsymbol{\alpha}_1\right)$，$\boldsymbol{e}_2,\cdots,\boldsymbol{e}_k$，在这组基底下，诸 $\boldsymbol{\alpha}_i$ 之坐标为 $\boldsymbol{\alpha}_1=\left(\dfrac{1}{\|\boldsymbol{\alpha}_1\|},0,\cdots,0\right)$，$\boldsymbol{\alpha}_2=(\xi_{21},\xi_{22},\cdots,\xi_{2k})$，$\cdots$，$\boldsymbol{\alpha}_{k+2}=(\xi_{k+2,1},\xi_{k+2,2},\cdots,\xi_{k+2,k})$，故由 $(\boldsymbol{\alpha}_i,\boldsymbol{\alpha}_j)<0$ 即得 $\xi_{i,1}<0$ $(i=2,\cdots,k+2)$，且有 $\sum\limits_{l=2}^{k}\xi_{il}\xi_{jl}<0$ $(i,j=2,\cdots,k+2;\ i\neq j)$. 而第二个式子就说明了 $k+1$ 个 $k-1$ 元向量 $(\xi_{22},\xi_{23},\cdots,\xi_{2k})$，……，$(\xi_{k+2,2},\xi_{k+2,3},\cdots,\xi_{k+2,k})$ 之数积为负数，这与归纳假定 $k-1$ 维空间能成立相矛盾.

练习三十三 **1**. 提示：因齐式之矩阵

$$A=\begin{pmatrix}1&0&2\sqrt{2}\\0&5&0\\2\sqrt{2}&0&-1\end{pmatrix}$$

之三个特征值为 5, 3, -3，故典型式为 $5x'^2+3y'^2-3z'^2$. 对应于 5, 3, -3 之 A 的特征向量分别可取为 $(0,1,0)$，$(\sqrt{2},0,1)$，$\left(-\dfrac{1}{\sqrt{2}},0,1\right)$；各个标准化后分别变为 $(0,1,0)$，$\left(\dfrac{\sqrt{2}}{\sqrt{3}},0,\dfrac{1}{\sqrt{3}}\right)$，$\left(-\dfrac{1}{\sqrt{3}},0,\dfrac{\sqrt{2}}{\sqrt{3}}\right)$，而有以之为列向量之直交矩阵 P，使得 $P^{-1}AP=\mathrm{diag}\{5,3-3\}$. **2**. 提示：因必有一组不全为 0 之 n 个实数 $x_1,\cdots,x_n$ 使 $AX=\lambda X$，$X'=(x_1,x_2,\cdots,x_n)$，于是，$\varphi(x_1,\cdots,x_n)=X'AX=\lambda X'X=\lambda(x_1^2+\cdots+x_n^2)$. **3**. 提示：有直交矩阵 P 使 $X=PY$，且 $X'AX=Y'P^{-1}APY=\lambda_1y_1^2+\cdots+\lambda_ny_n^2\leqslant\lambda(y_1^2+\cdots+y_n^2)$，且 $\geqslant u(y_1^2+\cdots+y_n^2)$，但 $y_1^2+\cdots+y_n^2=Y'Y=Y'EY=Y'P'PY=X'X$. **4**. 提示：$(\boldsymbol{x}A,\boldsymbol{y})=\left(\sum\limits_{i=1}^n a_{i1}x_i\right)y_1+\cdots+\left(\sum\limits_{i=1}^n a_{in}x_i\right)y_n=\sum\limits_{i,j=1}^n a_{ij}x_iy_j$，又 $(\boldsymbol{x},\boldsymbol{y}A)=\sum\limits_{i,j=1}^n a_{ij}y_ix_j$，再利用 $a_{ji}=a_{ij}$. **5**. 提示：$m(\lambda)$ 无重根 $\Rightarrow m(\lambda)=\lambda\left[\lambda^2-n^2\lambda-\dfrac{1}{12}(n-1)n^2(n+1)\right]$ ——参看第五章.

练习三十四 **1**. 提示：因有旋转变换

$$\begin{cases} x = x'\cos\theta - y'\sin\theta, \\ y = x'\sin\theta + y'\cos\theta \end{cases}$$

使 $ax^2+2bxy+cy^2=\lambda_1 x'^2+\lambda_2 y'^2$. 故(19)变为

$$\lambda_1 x^2+\lambda_2 y^2+2d_1 x+2e_1 y+f=0$$

形. 分二种情况: $r_B=2$ 及 $r_B=1$. 当 $r_B=2$ 时, $\lambda_1\lambda_2\neq 0$, 再平移到新原点为 $\left(-\frac{d_1}{\lambda_1}, -\frac{e_1}{\lambda_2}\right)$, 就简化成 $\lambda_1 x^2+\lambda_2 y^2+f_1=0$ 形. 若 $r_A=3$, 则 $f_1\neq 0$ 且 $f_1=\frac{|A|}{|B|}$,故为双曲线($|B|=\lambda_1\lambda_2<0$ 时), 或椭圆($|B|>0$ 且 λ_1、λ_2 均与 $\frac{|A|}{|B|}$ 反号时, 其等价条件是 $(\lambda_1+\lambda_2)\cdot|A|=(a+c)\cdot|A|<0$ 与 $|B|>0$), 或虚椭圆(在 $(a+c)\cdot|A|>0$ 与 $|B|>0$ 时). 若 $r_A=2$, 则 $f_1=0$, 故为相交二直线(在 $\lambda_1\lambda_2=|B|<0$ 时)或一点(在 $|B|>0$ 时). 当 $r_B=1$ 时, 不妨令 $\lambda_1\neq 0$, $\lambda_2=0$, 这时再平移可得 $\lambda_1 x^2+2e_1 y+f_1=0$ 形. 故若 $r_A=3$, 则 $e_1\neq 0$, 曲线为抛物线; 若 $r_A=2$, 则 $e_1=0$ 而 $f\neq 0$, 曲线为二平行直线(实或虚); 若 $r_A=1$, 则 $e_1=f_1=0$, 曲线为二个重合的直线. **2**. 提示: i) 题中

$$A=\begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 5 \\ 3 & 4 & 5 & 6 \\ 4 & 5 & 6 & 7 \end{pmatrix}$$

之秩 $r_A=2$, $r_B=2$, B 之特征值为

$$\lambda_1=\frac{9+\sqrt{105}}{2},\quad \lambda_2=\frac{9-\sqrt{105}}{2},\quad \lambda_3=0,$$

于是经适当的轴旋转变换, 方程可简化为 $\lambda_1 x^2+\lambda_2 y^2+2gx+2hy+2kz+7=0$, 因而再适当地平移可简化为 $\lambda_1 x^2+\lambda_2 y^2+2kz+l=0$, 故 $r_A=2\Rightarrow k=l=0$, 又由于 $\lambda_1\lambda_2<0$ 知为二个相交的平面. ii) 这时 B 之三个特征值为 1, 3, -5, 故方程可简化为 $x^2+3y^2-5z^2+l=0$ 形; 又由于 $-2=|A|=1\cdot 3\cdot(-5)\cdot l=-15l$, 知 $l=\frac{2}{15}$, 故为双叶双曲面.

练习三十五 **1**. 提示: 取 n 维空间 $\mathfrak{R}_n$ 之一组基底 $e_1, \cdots, e_n$, 令 $(e_i, e_j)=a_{ji}$ 得使 $\mathfrak{R}_n$ 为酉空间, 于是柯西不等式即为上不等式. **2**. 提示: 取 $x=e_i$, $y=e_j$ 即得 $(e_i, e_j)=\delta_{ij}$.

练习三十六 **1**. 提示: 酉矩阵 $A=(a_{ij})$ 之第 k 行 $a_{k1},\cdots,a_{kn}$ 用 ζ 去乘(但 $|\zeta|=1$)后的矩阵令为 $B=(b_{ij})$, 则 $B=E_k(\zeta)\cdot A$, 而 $E_k(\zeta)$ 显为酉矩阵. **2**. 提示: i) 与 ii) $\Rightarrow \bar{A}'=A^{-1}$、$\bar{A}'=A\Rightarrow A^{-1}=A\Rightarrow A^2=E\Rightarrow$ iii). i) 与 iii) $\Rightarrow \bar{A}'=A^{-1}$、$A^2=E\Rightarrow A=A^{-1}=\bar{A}'\Rightarrow$ ii). ii) 与 iii) $\Rightarrow \bar{A}'=A$、$A^2=E\Rightarrow A^{-1}=A=\bar{A}'\Rightarrow$ i).

3. 提示: i)中 $\boldsymbol{A}$ 之三个特征值为 0, 0, 1. 对应于 1 之 $\boldsymbol{A}$ 之特征单位向量为

$$\boldsymbol{X}_1=\begin{pmatrix}-\frac{i}{\sqrt{3}}\\ \frac{i}{\sqrt{6}}\\ \frac{1}{\sqrt{2}}\end{pmatrix};$$

对应于 0 之 $\boldsymbol{A}$ 之两个互为直交且长为 1 之特征向量为

$$\boldsymbol{X}_2=\begin{pmatrix}\frac{1}{\sqrt{3}}\\ \frac{\sqrt{2}}{\sqrt{3}}\\ 0\end{pmatrix} \quad 与 \quad \boldsymbol{X}_3=\begin{pmatrix}\frac{1}{\sqrt{3}}i\\ -\frac{1}{\sqrt{6}}i\\ \frac{1}{\sqrt{2}}\end{pmatrix}.$$

于是酉矩阵 $\boldsymbol{U}=(\boldsymbol{X}_1\boldsymbol{X}_2\boldsymbol{X}_3)$ 具性质

$$\boldsymbol{U}^{-1}\boldsymbol{A}\boldsymbol{U}=\begin{pmatrix}1&&\\&0&\\&&0\end{pmatrix}.$$

又 ii)中 $\boldsymbol{A}$ 之三个特征值为 0, $\pm\sqrt{2}$; 对应于 $\sqrt{2}$, $-\sqrt{2}$, 0 的特征单位向量各可取为

$$\boldsymbol{X}_1=\begin{pmatrix}\frac{1}{\sqrt{2}}\\ -\frac{1}{2}i\\ \frac{1}{2}\end{pmatrix},\ \boldsymbol{X}_2=\begin{pmatrix}\frac{1}{\sqrt{2}}\\ \frac{1}{2}i\\ -\frac{1}{2}\end{pmatrix},\ \boldsymbol{X}_3=\begin{pmatrix}0\\ \frac{i}{\sqrt{2}}\\ \frac{1}{\sqrt{2}}\end{pmatrix},$$

于是酉矩阵 $\boldsymbol{U}=(\boldsymbol{X}_1, \boldsymbol{X}_2, \boldsymbol{X}_3)$ 使 $\boldsymbol{U}^{-1}\boldsymbol{A}\boldsymbol{U}=\mathrm{diag}\{\sqrt{2}, -\sqrt{2}, 0\}$. 4. 提示: 设 $\boldsymbol{A}$ 之 n 个特征值为 $\lambda_1, \cdots, \lambda_n$(故为实数). 令 $k>\max(|\lambda_1|, \cdots, |\lambda_n|)$, 则每 $\lambda_i>-k$. 若 λ_0 为 $\boldsymbol{B}=\mu\boldsymbol{E}+\boldsymbol{A}$ 之任一特征值, 则 $0=|\lambda_0\boldsymbol{E}-(\mu\boldsymbol{E}+\boldsymbol{A})|$, 故 $\lambda_0-\mu=$某 $\lambda_i>-k$, 所以 $\lambda_0>\mu-k\geqslant 0$. 5. 提示: α 为 $\boldsymbol{A}(\overline{\boldsymbol{A}}'=-\boldsymbol{A})$之特征值, 有特征向量 $\boldsymbol{X}\neq 0$ 使 $\boldsymbol{A}\boldsymbol{X}=\alpha\boldsymbol{X}\Rightarrow\bar{\alpha}\overline{\boldsymbol{X}}'\boldsymbol{X}=\overline{\boldsymbol{X}}'\overline{\boldsymbol{A}}'\boldsymbol{X}=-\overline{\boldsymbol{X}}'\boldsymbol{A}\boldsymbol{X}=-\alpha\overline{\boldsymbol{X}}'\boldsymbol{X}\Rightarrow(\alpha+\bar{\alpha})\overline{\boldsymbol{X}}'\boldsymbol{X}\Rightarrow\alpha=-\bar{\alpha}\Rightarrow\alpha$ 纯虚数或零. 又令 β 为 $\boldsymbol{A}$ 之另一特征值($\beta\neq\alpha$), $\boldsymbol{Y}$ 为其特征向量, $\boldsymbol{A}\boldsymbol{Y}=\beta\boldsymbol{Y}\Rightarrow\bar{\alpha}\overline{\boldsymbol{X}}'\boldsymbol{Y}=\overline{\boldsymbol{X}}'\overline{\boldsymbol{A}}'\boldsymbol{Y}=-\overline{\boldsymbol{X}}'\boldsymbol{A}\boldsymbol{Y}=-\overline{\boldsymbol{X}}'\beta\boldsymbol{Y}\Rightarrow(\bar{\alpha}+\beta)\overline{\boldsymbol{X}}'\boldsymbol{Y}=0$; 但 $\beta\neq\alpha=-\bar{\alpha}\Rightarrow\bar{\alpha}+\beta\neq 0\Rightarrow\overline{\boldsymbol{X}}'\boldsymbol{Y}=0$, 即对应于不同特征值之特征向量直交. 再取适合 $\boldsymbol{A}\boldsymbol{X}_1=\alpha\boldsymbol{X}_1$ 之单位向量 $\boldsymbol{X}_1$, 即 $\overline{\boldsymbol{X}}_1'\boldsymbol{X}_1=1$, 作酉矩阵 $\boldsymbol{U}_1=(\boldsymbol{X}_1, \cdots)$, 故

$U_1^{-1}AU_1$ 之第一个列向量为

$$U_1^{-1}AX_1=\alpha U_1^{-1}X_1=\begin{pmatrix}\alpha\\0\\\vdots\\0\end{pmatrix}\Rightarrow U_1^{-1}AU_1=\begin{pmatrix}\alpha & s_2 \cdots s_n\\0 & \\\vdots & B\\0 & \end{pmatrix}$$

之反厄米特性保证了 $s_2=\cdots=s_n=0$ 及 $\bar{B}'=-B$. 由归纳法有 $n-1$ 级酉矩阵 W 使 $W^{-1}BW=\mathrm{diag}\{\alpha_2, \cdots, \alpha_n\}$；再令 $U_2=\begin{pmatrix}1 & \\ & W\end{pmatrix}$，则 $\bar{U}_2'=U_2^{-1}$，故 $U=U_1U_2$ 为酉矩阵且有

$$U^{-1}AU=\begin{pmatrix}1 & \\ & W^{-1}\end{pmatrix}\begin{pmatrix}\alpha & \\ & B\end{pmatrix}\begin{pmatrix}1 & \\ & W\end{pmatrix}=\mathrm{diag}\{\alpha_1\alpha_2, \cdots, \alpha_n\}.$$

于是最小多项式无重根，且每特征值之重数等于其特征子空间的维数.

6. 提示：$E_{m+n}=\bar{P}'P=\begin{pmatrix}\bar{P}_1' & O\\ \bar{A}' & \bar{P}_2'\end{pmatrix}\begin{pmatrix}P_1 & A\\ O & P_2\end{pmatrix}=\begin{pmatrix}\bar{P}_1'P_1 & \bar{P}_1'A\\ \bar{A}'P_1 & \bar{A}'A+\bar{P}_2'P_2\end{pmatrix}\Rightarrow \bar{P}_1'P_1=E_m\Rightarrow P_1$ 酉矩阵；故再由 $\bar{P}_1'A=O$ 知 $A=O\Rightarrow\bar{P}_2'P_2=E_n\Rightarrow P_2$ 酉矩阵.

7. 提示：易知 $\bar{A}'A=\bar{B}'B$ 为正定的，故有满秩阵 Q 使 $\bar{Q}'\bar{A}'AQ=\bar{Q}'\bar{B}'BQ=E$，即 $(\overline{AQ})'(AQ)=(\overline{BQ})'(BQ)=E$，因之 $AQ=V_1$ 与 $BQ=V_2$ 均为酉矩阵，故 $A=V_1Q^{-1}=V_1V_2^{-1}B=UB$，$U=V_1V_2^{-1}$. 8. 提示：$S$ 之实反对称$\Rightarrow$特征值为 0 或纯虚数(问题 5)$\Rightarrow E-S$ 与 $E+S$ 是满秩的. 又显有

$$A=(E-S)(E+S)^{-1}=(E+S)^{-1}(E-S),$$

故

$$\begin{aligned}A'A&=(E-S)'\cdot[(E+S)^{-1}]'\cdot(E-S)\cdot(E+S)^{-1}\\&=(E-S')[(E+S)']^{-1}(E-S)(E+S)^{-1}\\&=(E+S)(E-S)^{-1}(E-S)(E+S)^{-1}=E.\end{aligned}$$

9. 提示：$E-S$ 与 $E+S$ 仍为满秩，且

$$\begin{aligned}\bar{A}'A&=(\overline{E-S})'[(\overline{E+S})']^{-1}(E-S)(E+S)^{-1}\\&=(E+S)(E-S)^{-1}(E-S)(E+S)^{-1}=E\end{aligned}$$

说明了 $(E-S)\cdot(E+S)^{-1}$ 是酉交的.

练习三十七 1. 提示：有实直交矩阵 T 使

$$T^{-1}AT=\begin{pmatrix}1 & 0 & 0\\0 & \cos\theta & -\sin\theta\\0 & \sin\theta & \cos\theta\end{pmatrix}=B,$$

故 $|\lambda E-A|=(\lambda-1)(\lambda^2-2\cos\theta\cdot\lambda+1)=\lambda^3-t\lambda^2+t\lambda-1$，$t=1+2\cos\theta$，故 $-1\leqslant t\leqslant 3$. 2. 提示：A 与 B 中一为正常的，一为非正常的，于是 $A^{-1}B$ 为非正常的，因而有 -1 为其特征值，于是 $E+A^{-1}B$ 降秩，故 $A(E+A^{-1}B)=A+B$ 降秩. 3. 提示：$\bar{U}'U=(P'-iQ')(P+iQ)=(P'P+Q'Q)+i(P'Q-Q'P)$,

练习三十八　1. 提示：在本节定理之证明中，$\boldsymbol{Q}$ 是实的，$\boldsymbol{A}_1$ 实对称，$\boldsymbol{U}_1$ 得取为实直交的．　2. 提示：由本节的定理，已知 $|\lambda\boldsymbol{B}-\boldsymbol{A}|=0$ 之根是实数，且为 $\boldsymbol{A}_1=\overline{\boldsymbol{Q}}'\boldsymbol{A}\boldsymbol{Q}$ 之特征值；故从 $\boldsymbol{A}_1$ 之正定性即得．　3. 提示：对角矩阵、酉矩阵、厄米特及反厄米特矩阵 $\boldsymbol{A}$ 显然都满足 $\overline{\boldsymbol{A}}'\boldsymbol{A}=\boldsymbol{A}\overline{\boldsymbol{A}}'$ 的要求．又

$$
\begin{aligned}
(\overline{\overline{\boldsymbol{U}}'\boldsymbol{A}\boldsymbol{U}})'\cdot(\overline{\boldsymbol{U}}'\boldsymbol{A}\boldsymbol{U})&=\overline{\boldsymbol{U}}'\overline{\boldsymbol{A}}'\boldsymbol{U}\cdot\overline{\boldsymbol{U}}'\boldsymbol{A}\boldsymbol{U}=\overline{\boldsymbol{U}}'\overline{\boldsymbol{A}}'\boldsymbol{E}\boldsymbol{A}\boldsymbol{U}=\overline{\boldsymbol{U}}'\overline{\boldsymbol{A}}'\boldsymbol{A}\boldsymbol{U}\\
&=\overline{\boldsymbol{U}}'\boldsymbol{A}\overline{\boldsymbol{A}}'\boldsymbol{U}=\overline{\boldsymbol{U}}'\boldsymbol{A}\boldsymbol{E}\overline{\boldsymbol{A}}'\boldsymbol{U}=\overline{\boldsymbol{U}}'\boldsymbol{A}\boldsymbol{U}\overline{\boldsymbol{U}}'\overline{\boldsymbol{A}}'\boldsymbol{U}\\
&=(\overline{\boldsymbol{U}}'\boldsymbol{A}\boldsymbol{U})\cdot(\overline{\overline{\boldsymbol{U}}'\boldsymbol{A}\boldsymbol{U}})'
\end{aligned}
$$

即说明了 $\overline{\boldsymbol{U}}'\boldsymbol{A}\boldsymbol{U}$ 之规范性；　4. 提示：$\boldsymbol{U}^{-1}\boldsymbol{A}\boldsymbol{U}=\boldsymbol{D}$ 之对角性 $\Rightarrow\boldsymbol{D}$ 之规范性，$\boldsymbol{A}=\boldsymbol{U}\boldsymbol{D}\boldsymbol{U}^{-1}=(\overline{\boldsymbol{U}^{-1}})'\boldsymbol{D}(\boldsymbol{U}^{-1})\Rightarrow\boldsymbol{A}$ 之规范性（第 3 题）．反之，$\boldsymbol{A}$ 之规范性 $\Rightarrow$ $\overline{\boldsymbol{A}}'\boldsymbol{A}=\boldsymbol{A}\overline{\boldsymbol{A}}'$；取 λ_1 为 $\boldsymbol{A}$ 之一特征值，$\boldsymbol{X}_1$ 为相应之单位特征向量，$\boldsymbol{U}_1=(\boldsymbol{X}_1,\ \cdots)$ 为以 $\boldsymbol{X}_1$ 为第一个列向量的酉矩阵，则从 $\boldsymbol{A}\boldsymbol{X}_1=\lambda_1\boldsymbol{X}_1$ 得

$$
\boldsymbol{A}_1=\boldsymbol{U}_1^{-1}\boldsymbol{A}\boldsymbol{U}_1=\begin{pmatrix}\lambda_1 & s_2\ \cdots\ s_n\\ 0 & \\ \vdots & \boldsymbol{B}\\ 0 & \end{pmatrix},
$$

于是 $\boldsymbol{A}_1$ 之规范性（上题）$\Rightarrow\overline{\boldsymbol{A}}_1'\boldsymbol{A}_1=\boldsymbol{A}_1\overline{\boldsymbol{A}}_1'\Rightarrow\bar{\lambda}_1\lambda_1=\bar{\lambda}_1\lambda_1+\bar{s}_2s_2+\cdots+\bar{s}_ns_n\Rightarrow s_2=\cdots=s_n=0\Rightarrow\overline{\boldsymbol{B}}'\boldsymbol{B}=\boldsymbol{B}\overline{\boldsymbol{B}}'$；用归纳法，有 $n-1$ 级酉矩阵 $\boldsymbol{W}$ 使

$$
\boldsymbol{W}^{-1}\boldsymbol{B}\boldsymbol{W}=\operatorname{diag}\{\lambda_2,\ \cdots,\ \lambda_n\},
$$

再令　$\boldsymbol{U}_2=\begin{pmatrix}1 & \\ & \boldsymbol{W}\end{pmatrix}$，　$\boldsymbol{U}=\boldsymbol{U}_1\boldsymbol{U}_2$，则

$$
\begin{aligned}
\boldsymbol{U}^{-1}\boldsymbol{A}\boldsymbol{U}=\boldsymbol{U}_2^{-1}\begin{pmatrix}\lambda_1 & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{B}\end{pmatrix}\boldsymbol{U}_2=\begin{pmatrix}\lambda_1 & \\ & \boldsymbol{W}^{-1}\boldsymbol{B}\boldsymbol{W}\end{pmatrix}\\
=\operatorname{diag}\{\lambda_1,\ \lambda_2,\ \cdots,\ \lambda_n\}.
\end{aligned}
$$

5. 提示：由上题有酉矩阵 $\boldsymbol{U}$ 使 $\boldsymbol{U}^{-1}\boldsymbol{A}\boldsymbol{U}=\operatorname{diag}\{\lambda_1,\ \lambda_2,\ \cdots,\ \lambda_n\}=\boldsymbol{D}$．每 λ_i 实 $\Rightarrow\overline{\boldsymbol{D}}'=\boldsymbol{D}\Rightarrow\overline{\boldsymbol{A}}'=\boldsymbol{A}$，即 i) 成立．每 λ_i 为 0 或纯虚数 $\Rightarrow\overline{\boldsymbol{D}}'=-\boldsymbol{D}\Rightarrow\overline{\boldsymbol{A}}'=-\boldsymbol{A}$，即 ii) 成立．每 $|\lambda_i|=1\Rightarrow\overline{\boldsymbol{D}}'\boldsymbol{D}=\boldsymbol{E}\Rightarrow\overline{\boldsymbol{A}}'\boldsymbol{A}=\boldsymbol{E}$，即 iii) 成立．　6. 提示：存在酉矩阵 $\boldsymbol{U}$，使 $\boldsymbol{U}^{-1}\boldsymbol{A}\boldsymbol{U}=\operatorname{diag}\{\lambda_1,\ \cdots,\ \lambda_n\}=\boldsymbol{D}\Rightarrow\boldsymbol{U}^{-1}\overline{\boldsymbol{A}}'\boldsymbol{A}\boldsymbol{U}=\operatorname{diag}\{\bar{\lambda}_1\lambda_1,\ \cdots,\ \bar{\lambda}_n\lambda_n\}\Rightarrow$ $\bar{\lambda}_1\lambda_1,\ \cdots,\ \bar{\lambda}_n\lambda_n$ 与 $\mu_1,\ \cdots,\ \mu_n$ 一致 $\Rightarrow\sum\mu_i=\sum|\lambda_i|^2$．　7. 提示：只需证条件的必要性．设 $\boldsymbol{A}=(a_{ij})$ 为上三角阵即 $a_{ij}=0\,(i>j$ 时)．若 $\overline{\boldsymbol{A}}'\boldsymbol{A}=\boldsymbol{A}\overline{\boldsymbol{A}}'$，则考虑 $\overline{\boldsymbol{A}}'\boldsymbol{A}=\boldsymbol{A}\overline{\boldsymbol{A}}'$ 之第一行第一列交叉处之元素有 $\bar{a}_{11}a_{11}=a_{11}\bar{a}_{11}+a_{12}\bar{a}_{12}+\cdots+a_{1n}\bar{a}_{1n}\Rightarrow a_{12}=\cdots=a_{1n}=0$；再归纳地假定 $a_{23}=\cdots=a_{2n}=0,\ \cdots,\ a_{i-1,i}=\cdots=a_{i-1,n}=0$，而再考虑 $\overline{\boldsymbol{A}}'\boldsymbol{A}=\boldsymbol{A}\overline{\boldsymbol{A}}'$ 之第 i 行第 i 列处之元，就有

$$
\bar{a}_{ii}a_{ii}=a_{ii}\bar{a}_{ii}+a_{i,i+1}\bar{a}_{i,i+1}+\cdots+a_{in}\bar{a}_{in}\Rightarrow a_{i,i+1}=\cdots=a_{in}=0.
$$

故归纳地证明了 $a_{st}=0\,(s<t)\to\boldsymbol{A}$ 为对角的．　8. 提示：取 $\boldsymbol{A}$ 之一特征值 λ_1 及其相应之单位特征向量 $\boldsymbol{X}_1$（即 $\boldsymbol{A}\boldsymbol{X}_1=\lambda_1\boldsymbol{X}_1$，$\overline{\boldsymbol{X}}_1'\boldsymbol{X}_1=1$），再作以 $\boldsymbol{X}_1$ 为第一

个列向量的酉矩阵

$$U_1=(X_1,\ \cdots)\Rightarrow U_1^{-1}AU_1=\begin{pmatrix}\lambda_1 & * \\ 0 & B\end{pmatrix}.$$

B 为 $n-1$ 级矩阵，由归纳假设，有 $n-1$ 级酉矩阵 V 使 $V^{-1}BV=\mathrm{diag}\{b_1,\ \cdots,\ b_{n-1}\}$；再令 $U_2=\begin{pmatrix}1 & \\ & V\end{pmatrix}$，则令 $U=U_1U_2$ 后有

$$U^{-1}AU=U_2^{-1}\begin{pmatrix}\lambda_1 & * \\ 0 & B\end{pmatrix}U_2=\begin{pmatrix}1 & \\ & V^{-1}\end{pmatrix}\begin{pmatrix}\lambda_1 & * \\ 0 & B\end{pmatrix}\begin{pmatrix}1 & \\ & V\end{pmatrix}$$

$$=\begin{pmatrix}\lambda_1 & *V \\ 0 & V^{-1}BV\end{pmatrix}=\begin{pmatrix}\lambda_1 & & * \\ & b_1 & \\ & & \ddots \\ & & & b_{n-1}\end{pmatrix}.$$

若先考虑 $\overline{A}'$，则有酉矩阵 W 使 $W^{-1}\overline{A}'W$ 为上三角矩阵，故 $W^{-1}AW$ 为下三角阵。 9．提示：两个二次齐式分别为$X'AX$，$X'BX$，$A=\begin{pmatrix}0 & -1 & 0 \\ -1 & -1 & 1 \\ 0 & 1 & 0\end{pmatrix}$，$B=\begin{pmatrix}2 & 1 & -2 \\ 1 & 2 & -2 \\ -2 & -2 & 3\end{pmatrix}$，$B$ 正定．有实满秩阵 Q 使 $Q'BQ=E$，由第六章可具体求一个

$$Q=\begin{pmatrix}\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} & \frac{2}{\sqrt{3}} \\ 0 & \frac{\sqrt{2}}{3} & \frac{2}{\sqrt{3}} \\ 0 & 0 & \sqrt{3}\end{pmatrix}.$$

由是对称阵 $$A_1=Q'AQ=\begin{pmatrix}0 & -\frac{1}{\sqrt{3}} & -\frac{\sqrt{2}}{\sqrt{3}} \\ -\frac{1}{\sqrt{3}} & 0 & 0 \\ -\frac{\sqrt{2}}{\sqrt{3}} & 0 & 0\end{pmatrix},$$

其特征值为 1，-1，0，分别对应于它们的单位特征向量为实直交矩阵

$$U_1=\begin{pmatrix}\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{\sqrt{2}}{\sqrt{3}} \\ -\frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{3}}\end{pmatrix}$$

之第一、二、三列，故

$$U_1^{-1}A_1U_1=\begin{pmatrix}1&&\\&-1&\\&&0\end{pmatrix}.$$

$$P=QU_1=\begin{pmatrix}0&1&-1\\-1&1&0\\-1&1&-1\end{pmatrix}$$

为所需之满秩变换，使典型式分别为 x^2-y^2 及 $x^2+y^2+z^2$.　　**10.** 提示：有非零向量 $X\neq 0$ 使 $UAX=\omega X$，故

$$|\omega|^2\bar{X}'X=\bar{\omega}\omega\bar{X}'X=\bar{X}'\bar{A}'AX=\bar{X}'\begin{pmatrix}|a_1|^2&&0\\&\ddots&\\0&&|a_n|^2\end{pmatrix}\cdot X\leqslant M^2\bar{X}'X,$$

而 $\geqslant m^2\bar{X}'X$.　　**11.** 提示：A 为酉矩阵与 $|A-\omega B|=0$ 导致 $\omega\neq 0$，由是有 $|\omega^{-1}E-\bar{A}'B|=0$. 又 B 之规范性说明有一酉矩阵 Q，使

$$Q^{-1}BQ=\begin{pmatrix}a_1&&0\\&\ddots&\\0&&a_n\end{pmatrix},$$

而 $Q^{-1}\bar{A}'Q$ 仍为酉矩阵，且 ω^{-1} 为 $Q^{-1}\bar{A}'Q\cdot Q^{-1}BQ=Q^{-1}\bar{A}'BQ$ 之一特征值，故由前题即得本题之正确性.

练习三十九　**1.** 提示：利用(9)式即知

$$\sin^2 A+\cos^2 A=\frac{1}{2}(e^{iA}e^{-iA}+e^{-iA}e^{iA})=E;$$

又因 A 与 $2\pi iE$ 相乘可交换，故

$$e^{A+2\pi iE}=e^Ae^{2\pi iE}=e^A\cdot\left\{\sum_{m=0}^{\infty}\frac{1}{m!}(2\pi i)^m\right\}E=e^A\cdot e^{2\pi i}E=e^A\cdot E;$$

同理，

$$\begin{aligned}\sin(A+2\pi E)&=\sin A\cdot\cos 2\pi E+\cos A\cdot\sin 2\pi E\\&=\sin A\cdot(\cos 2\pi)E+\cos A\cdot(\sin 2\pi)E=\sin A.\end{aligned}$$

2. 提示：从 $A'=-A(\bar{A}=A)$ 得

$$(e^A)^{-1}=e^{-A}=e^{A'}=E+A'+\frac{1}{2!}A'^2+\cdots=\left(E+A+\frac{1}{2!}A^2+\cdots\right)'=(e^A)',$$

即 e^A 实直交. 同理，从 $\bar{A}'=-A$ 得

$$(e^A)^{-1}=e^{-A}=e^{\bar{A}'}=\left(E+\bar{A}+\frac{1}{2!}\bar{A}^2+\cdots\right)'=(\overline{e^A})'$$

即 e^A 酉矩阵.　　**3.** 提示：从 $\bar{A}'=A$ 得 $(\overline{iA})'=-i\bar{A}'$ 说明 iA 之反厄米特性，故由上题即得. 在 $n=1$ 时，$\bar{A}'=A$ 说明 A 为一实数，故 $e^{iA}=\cos A+i\sin A$ 表示模为 1 之复数，即在单位圆上.　　**4.** 提示：反复利用定理 2，可知公式一般不成立. 只是在 $\dot{A}(t)\cdot A(t)=A(t)\cdot\dot{A}(t)$ 时，才成立(用归纳法证明).

5. 提示: $\boldsymbol{A}$ 之特征值为 $1, \omega, \omega^2$(ω 为 1 之虚立方根), 两两互异, 故有满秩阵 $\boldsymbol{P}$ 使 $\boldsymbol{P}^{-1}\boldsymbol{AP}=\mathrm{diag}\{1, \omega, \omega^2\}$, 于是仿例 2, 可知 $e\boldsymbol{A}=\boldsymbol{P}\cdot\mathrm{diag}\{e, e^{\omega}, e^{\omega^2}\}\cdot\boldsymbol{P}^{-1}$, 相应地有关于 $\sin\boldsymbol{A}$ 及 $\cos\boldsymbol{A}$ 之式子. 又 $\boldsymbol{P}$ 之三个列向量可分别令为对应于 $1,\omega,\omega^2$ 之特征向量, 故可取 $\boldsymbol{P}=\begin{pmatrix}1&1&1\\1&\omega&\omega^2\\1&\omega^2&\omega\end{pmatrix}$, 因之 $\boldsymbol{P}^{-1}=\dfrac{1}{3}\begin{pmatrix}1&1&1\\1&\omega^2&\omega\\1&\omega&\omega^2\end{pmatrix}$,

而有

$$e\boldsymbol{A}=\frac{1}{3}\begin{pmatrix}\alpha&\gamma&\beta\\\beta&\alpha&\gamma\\\gamma&\beta&\alpha\end{pmatrix},$$

$$\alpha=e+e^{\omega}+e^{\omega^2},\quad \beta=e+\omega e^{\omega}+\omega^2e^{\omega^2},\quad \gamma=e+\omega^2e^{\omega}+\omega e^{\omega^2}.$$

同理可求得 $\sin\boldsymbol{A}$ 及 $\cos\boldsymbol{A}$.

练习四十 1. 提示: 直接由定义自明. 2. 提示: $\overline{\boldsymbol{A}}'=-\boldsymbol{A}$ 及 $\overline{\boldsymbol{B}}'=-\boldsymbol{B}\Rightarrow(\overline{\boldsymbol{A}\times\boldsymbol{B}})'=(\overline{\boldsymbol{A}}\times\overline{\boldsymbol{B}})'=\overline{\boldsymbol{A}}'\times\overline{\boldsymbol{B}}'=(-\boldsymbol{A})\times(-\boldsymbol{B})=\boldsymbol{A}\times\boldsymbol{B}$. 对其他情况可类似地解决. 3. 提示: $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 分别为 n 级与 m 级的$\Rightarrow(\boldsymbol{A}\times\boldsymbol{B})^*=|\boldsymbol{A}\times\boldsymbol{B}|\cdot(\boldsymbol{A}\times\boldsymbol{B})^{-1}=|\boldsymbol{A}|^m|\boldsymbol{B}|^n\cdot(\boldsymbol{A}^{-1}\times\boldsymbol{B}^{-1})=|\boldsymbol{A}|^{m-1}\cdot|\boldsymbol{B}|^{n-1}\cdot(\boldsymbol{A}^*\times\boldsymbol{B}^*)$.

练习四十一 1. 提示: 利用定理 4 及其推论 1 得 i); 又由定理 2 与定理 4 可得 ii). 2. 提示: 因有 $\boldsymbol{P}$ 使 $\boldsymbol{P}^{-1}\boldsymbol{AP}=\mathrm{diag}\{\lambda_1, \lambda_2, \cdots, \lambda_n\}$, 故 $(\boldsymbol{P}^{-1}\boldsymbol{AP})^{(k)}=\mathrm{diag}\{p_1, p_2, \cdots, p_{\binom{n}{k}}\}$, 每 p_i 为 k 个 λ_j 的乘积, 即 $(\boldsymbol{P}^{(k)})^{-1}\cdot\boldsymbol{A}^{(k)}\cdot\boldsymbol{P}^{(k)}=\mathrm{diag}\{p_1, p_2, \cdots\}$, $\boldsymbol{A}^{(k)}$ 之最小多项式无重根. 3. 提示: $\boldsymbol{A}^{(2)}=\begin{pmatrix}1&0&0\\0&0&0\\0&0&0\end{pmatrix}$, $m_{\boldsymbol{A}^{(2)}}(\lambda)=\lambda(\lambda-1)$, $m_{\boldsymbol{A}}(\lambda)=\lambda(\lambda-1)^2$, 故上题之逆不成立. 4. 提示: 定理 3 保证了 $\boldsymbol{A}^{(k)}$ 是厄米特的. 因 $\boldsymbol{P}^{-1}\boldsymbol{AP}=\mathrm{diag}\{\lambda_1, \cdots, \lambda_n\}\Rightarrow(\boldsymbol{P}^{(k)})^{-1}\cdot\boldsymbol{A}^{(k)}\cdot\boldsymbol{P}^{(k)}=\mathrm{diag}\{p_1, \cdots, p_{\binom{n}{k}}\}$, 每 p_i 为 k 个 λ_j 之积. 当每 $\lambda_i>0$ 时, 每 $p_i>0$, 故 $\boldsymbol{A}^{(k)}$ 正定. 当每 $\lambda_i<0$ 时, $p_i>0$ 或 <0 全由 k 之偶或奇而定.

练习四十二 1. 提示: $\boldsymbol{P}^{-1}\boldsymbol{AP}=\boldsymbol{B}$(上三角, 即 $b_{ij}>0(i>j)$, 因之可令 $\lambda_i=b_{ii}$) $\Rightarrow\overline{\boldsymbol{B}}'\boldsymbol{B}=\boldsymbol{P}^{-1}\overline{\boldsymbol{A}}'\boldsymbol{AP}$($\boldsymbol{P}$ 可选为酉矩阵)$\Rightarrow\mathrm{tr}(\overline{\boldsymbol{B}}'\boldsymbol{B})=\mathrm{tr}(\overline{\boldsymbol{A}}'\boldsymbol{A})\Rightarrow\sum_{r=1}^{n}|\lambda_r|^2+\sum_{r<s}|b_{rs}|^2=\sum_{r,s=1}^{n}|a_{rs}|^2\Rightarrow$i) 成立. 若令 $\boldsymbol{B}_1=\dfrac{1}{2}(\boldsymbol{B}\pm\overline{\boldsymbol{B}}')$, $\boldsymbol{A}_1=\dfrac{1}{2}(\boldsymbol{A}\pm\overline{\boldsymbol{A}}')$, 则 $\boldsymbol{B}_1=\boldsymbol{P}^{-1}\boldsymbol{A}_1\boldsymbol{P}\Rightarrow\overline{\boldsymbol{B}}_1'\boldsymbol{B}_1=\boldsymbol{P}^{-1}\overline{\boldsymbol{A}}_1'\boldsymbol{A}_1\boldsymbol{P}\Rightarrow\mathrm{tr}(\overline{\boldsymbol{B}}_1'\boldsymbol{B}_1)=\mathrm{tr}(\overline{\boldsymbol{A}}_1'\boldsymbol{A}_1)\Rightarrow$可得 ii) 与 iii). 2. 提示: 令 $b_{ij}=a_{ij}/d_i(i, j=1, \cdots, n)$, 而作矩阵 $\boldsymbol{B}=(b_{ij})$, 则 $\det\boldsymbol{A}=d_1d_2\cdots d_n\cdot\det B$. 设 α 为 $\boldsymbol{B}$ 之一特征值, 则由定理 3 知有某 k 使 $|\alpha-b_{kk}|\leqslant\sum_{\substack{s=1\\s\neq k}}^{n}|b_{ks}|$, 因而 $|\alpha|\geqslant|b_{kk}|-\sum_{\substack{s=1\\s\neq k}}^{n}|b_{ks}|=1$, 于是 $\det B$ 之模 $\geqslant1$.

5. 提示: $\boldsymbol{A}$ 之特征值为 $1, \omega, \omega^2$(ω 为 1 之虚立方根), 两两互异, 故有满秩阵 $\boldsymbol{P}$ 使 $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}=\mathrm{diag}\{1, \omega, \omega^2\}$, 于是仿例 2, 可知 $e\boldsymbol{A}=\boldsymbol{P}\cdot\mathrm{diag}\{e, e^{\omega}, e^{\omega^2}\}\cdot\boldsymbol{P}^{-1}$, 相应地有关于 $\sin\boldsymbol{A}$ 及 $\cos\boldsymbol{A}$ 之式子. 又 $\boldsymbol{P}$ 之三个列向量可分别令为对应于 $1, \omega, \omega^2$ 之特征向量, 故可取 $\boldsymbol{P}=\begin{pmatrix}1 & 1 & 1\\ 1 & \omega & \omega^2\\ 1 & \omega^2 & \omega\end{pmatrix}$, 因之 $\boldsymbol{P}^{-1}=\frac{1}{3}\begin{pmatrix}1 & 1 & 1\\ 1 & \omega^2 & \omega\\ 1 & \omega & \omega^2\end{pmatrix}$,

而有
$$e\boldsymbol{A}=\frac{1}{3}\begin{pmatrix}\alpha & \gamma & \beta\\ \beta & \alpha & \gamma\\ \gamma & \beta & \alpha\end{pmatrix},$$
$$\alpha=e+e^{\omega}+e^{\omega^2},\quad \beta=e+\omega e^{\omega}+\omega^2 e^{\omega^2},\quad \gamma=e+\omega^2 e^{\omega}+\omega e^{\omega^2}.$$
同理可求得 $\sin\boldsymbol{A}$ 及 $\cos\boldsymbol{A}$.

练习四十 1. 提示: 直接由定义自明. 2. 提示: $\bar{\boldsymbol{A}}'=-\boldsymbol{A}$ 及 $\bar{\boldsymbol{B}}'=-\boldsymbol{B}\Rightarrow(\overline{\boldsymbol{A}\times\boldsymbol{B}})'=(\bar{\boldsymbol{A}}\times\bar{\boldsymbol{B}})'=\bar{\boldsymbol{A}}'\times\bar{\boldsymbol{B}}'=(-\boldsymbol{A})\times(-\boldsymbol{B})=\boldsymbol{A}\times\boldsymbol{B}$. 对其他情况可类似地解决. 3. 提示: $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 分别为 n 级与 m 级的 $\Rightarrow(\boldsymbol{A}\times\boldsymbol{B})^*=|\boldsymbol{A}\times\boldsymbol{B}|\cdot(\boldsymbol{A}\times\boldsymbol{B})^{-1}=|\boldsymbol{A}|^m|\boldsymbol{B}|^n\cdot(\boldsymbol{A}^{-1}\times\boldsymbol{B}^{-1})=|\boldsymbol{A}|^{m-1}\cdot|\boldsymbol{B}|^{n-1}\cdot(\boldsymbol{A}^*\times\boldsymbol{B}^*)$.

练习四十一 1. 提示: 利用定理 4 及其推论 1 得 i); 又由定理 2 与定理 4 可得 ii). 2. 提示: 因有 $\boldsymbol{P}$ 使 $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}=\mathrm{diag}\{\lambda_1, \lambda_2, \cdots, \lambda_n\}$, 故 $(\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P})^{(k)}=\mathrm{diag}\{p_1, p_2, \cdots, p_{\binom{n}{k}}\}$, 每 p_i 为 k 个 λ_j 的乘积, 即 $(\boldsymbol{P}^{(k)})^{-1}\cdot\boldsymbol{A}^{(k)}\cdot\boldsymbol{P}^{(k)}=\mathrm{diag}\{p_1, p_2, \cdots\}$, $\boldsymbol{A}^{(k)}$ 之最小多项式无重根. 3. 提示: $\boldsymbol{A}^{(2)}=\begin{pmatrix}1 & 0 & 0\\ 0 & 0 & 0\\ 0 & 0 & 0\end{pmatrix}$, $m_{\boldsymbol{A}^{(2)}}(\lambda)=\lambda(\lambda-1)$, $m_{\boldsymbol{A}}(\lambda)=\lambda(\lambda-1)^2$, 故上题之逆不成立. 4. 提示: 定理 3 保证了 $\boldsymbol{A}^{(k)}$ 是厄米特的. 因 $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}=\mathrm{diag}\{\lambda_1, \cdots, \lambda_n\}\Rightarrow(\boldsymbol{P}^{(k)})^{-1}\cdot\boldsymbol{A}^{(k)}\cdot\boldsymbol{P}^{(k)}=\mathrm{diag}\{p_1, \cdots, p_{\binom{n}{k}}\}$, 每 p_i 为 k 个 λ_j 之积. 当每 $\lambda_i>0$ 时, 每 $p_i>0$, 故 $\boldsymbol{A}^{(k)}$ 正定. 当每 $\lambda_i<0$ 时, $p_i>0$ 或 <0 全由 k 之偶或奇而定.

练习四十二 1. 提示: $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}=\boldsymbol{B}$(上三角, 即 $b_{ij}>0(i>j)$, 因之可令 $\lambda_i=b_{ii}$) $\Rightarrow\bar{\boldsymbol{B}}'\boldsymbol{B}=\boldsymbol{P}^{-1}\bar{\boldsymbol{A}}'\boldsymbol{A}\boldsymbol{P}$($\boldsymbol{P}$ 可选为酉矩阵)$\Rightarrow\mathrm{tr}(\bar{\boldsymbol{B}}'\boldsymbol{B})=\mathrm{tr}(\bar{\boldsymbol{A}}'\boldsymbol{A})\Rightarrow\sum_{r=1}^{n}|\lambda_r|^2+\sum_{r<s}|b_{rs}|^2=\sum_{r,s=1}^{n}|a_{rs}|^2\Rightarrow$ i) 成立. 若令 $\boldsymbol{B}_1=\frac{1}{2}(\boldsymbol{B}\pm\bar{\boldsymbol{B}}')$, $\boldsymbol{A}_1=\frac{1}{2}(\boldsymbol{A}\pm\bar{\boldsymbol{A}}')$, 则 $\boldsymbol{B}_1=\boldsymbol{P}^{-1}\boldsymbol{A}_1\boldsymbol{P}\Rightarrow\bar{\boldsymbol{B}}_1'\boldsymbol{B}_1=\boldsymbol{P}^{-1}\bar{\boldsymbol{A}}_1'\boldsymbol{A}_1\boldsymbol{P}\Rightarrow\mathrm{tr}(\bar{\boldsymbol{B}}_1'\boldsymbol{B}_1)=\mathrm{tr}(\bar{\boldsymbol{A}}_1'\boldsymbol{A}_1)\Rightarrow$ 可得 ii) 与 iii). 2. 提示: 令 $b_{ij}=a_{ij}/d_i(i, j=1, \cdots, n)$, 而作矩阵 $\boldsymbol{B}=(b_{ij})$, 则 $\det\boldsymbol{A}=d_1d_2\cdots d_n\cdot\det B$. 设 α 为 $\boldsymbol{B}$ 之一特征值, 则由定理 3 知有某 k 使 $|\alpha-b_{kk}|\leqslant\sum_{\substack{s=1\\ s\neq k}}^{n}|b_{ks}|$, 因而 $|\alpha|\geqslant|b_{kk}|-\sum_{\substack{s=1\\ s\neq k}}^{n}|b_{ks}|=1$, 于是 $\det\boldsymbol{B}$ 之模 $\geqslant1$.